

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:22:19 ; Search time 163 Seconds
(without alignments)
1100.962 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLAULLAAGVLACGS.....YKFYDNKRIDSTSGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001s.*

5: Geneseq2002s.*

6: Geneseq2003as.*

7: Geneseq2003bs.*

8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2477	100.0	464	3 AAY79993	Aay79993 Plasmid L
2	1920	77.5	364	2 AAR20108	Aar20108 Protein D
3	1251.5	50.5	358	6 ABB84614	Abb84614 E. coli g
4	1251.5	50.5	359	7 ABO61848	Ab061848 Klebsiell
5	1219.5	49.2	364	7 ADP05718	Adf05718 Bacterial
6	1210.5	48.9	359	6 ABM68758	Abm68758 Phototrab
7	1074	43.4	356	2 AAY52773	Aay52773 Treponema
8	661.5	26.7	451	2 AAY06589	Aay06589 Lipoprote
9	661.5	26.7	451	6 ABU04447	Abu04447 Human exp
10	651	26.3	144	2 AAY02348	Aay02348 Protein D
11	651	26.3	446	2 AAY06590	Aay06590 Lipoprote
12	606	24.5	324	2 AAY02352	Aay02352 A represe
13	606	24.5	324	4 AAG63234	Aag63234 Amino aci
14	606	24.5	324	7 ADB67693	Adb67693 HIV-1 Lip
15	597	24.1	326	2 AAY02354	Aay02354 A represe
16	597	24.1	326	4 AAG63236	Aag63236 Amino aci
17	597	24.1	326	7 ADB67697	Adb67697 HIV-1 Pro
18	594.5	24.0	411	4 AAG63235	Aag63235 Amino aci
19	594.5	24.0	411	7 ADB67695	Adb67695 HIV-1 Lip
20	588.5	23.8	411	2 AAY02353	Aay02353 A represe
21	585.5	23.6	413	2 AAY02355	Aay02355 A represe
22	585.5	23.6	413	4 AAG63237	Aag63237 Amino aci
23	585.5	23.6	413	7 ADB67699	Adb67699 HIV-1 Pro
24	565	22.8	220	2 AAY25375	Aay25375 HPV fusio
25	565	22.8	220	2 AAY25378	Aay25378 HPV fusio

26	565	22.8	220	2 AAY02631	Aay02631 Prot.D1/3
27	565	22.8	220	2 AAY02634	Aay02634 Prot.D1/3
28	565	22.8	421	5 AAM50663	Aam50663 Thioresox
29	563	22.7	278	2 AAY25385	Aay25385 HPV fusio
30	563	22.7	278	2 AAY02641	Aay02641 Prot. D1/3
31	561.5	22.7	227	2 AAY25384	Aay25384 HPV fusio
32	561.5	22.7	227	2 AAY02640	Aay02640 Prot. D1/3
33	559.5	22.6	273	2 AAY25376	Aay25376 HPV fusio
34	559.5	22.6	273	2 AAY02632	Aay02632 Prot.D1/3
35	559.5	22.6	371	2 AAY25377	Aay25377 HPV fusio
36	559.5	22.6	371	2 AAY02633	Aay02633 Prot. D1/3
37	558	22.5	220	8 ADL64005	Adl64005 Chimeric
38	558	22.5	220	8 ADL63966	Adl63966 Chimeric
39	558	22.5	220	8 ADL97935	Adl97935 Protein D
40	558	22.5	227	2 AAY25382	Aay25382 HPV fusio
41	558	22.5	227	2 AAY02638	Aay02638 Prot. D1/3
42	558	22.5	383	2 AAY25386	Aay25386 HPV fusio
43	558	22.5	383	2 AAY02642	Aay02642 Prot. D1/3
44	429.5	17.3	797	4 ABG29533	Abg29533 Novel hum
45	381.5	15.4	386	6 ADA34227	Ada34227 Acinetoba

ALIGNMENTS

RESULT 1

AAY79993

ID AAY79993 standard; protein; 464 AA.

AC AAY79993;

DT 15-MAY-2000 (first entry)

DE Plasmid LPD-LB1-III protein sequence.

KW Vaccine; non-typeable Haemophilus influenzae; ntH1; infection;

KW chimeric protein; Haemophilus influenzae; P5-like fimbria protein;

KW lipoprotein D; LB1(f); immunogenic; antigenic; otitis media; sinusitis;

KW conjunctivitis; lower respiratory tract infection.

OS Haemophilus influenzae.

OS Synthetic.

XX WO9964067-A2.

XX 16-DEC-1999.

XX 28-MAY-1999; 99WO-US011980.

XX 11-JUN-1998; 98GB-00012613.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX (OHIS) UNIV OHIO STATE RES FOUND.

XX Bakaletz LO, Cohen J, Dequesne G, Lobet Y;

XX WPI; 2000-116457/10.

XX N-PSDB; AAZ91252.

XX Novel antigenic P5-like fimbria subunit peptides used in vaccines against

XX Haemophilus influenzae.

XX Claim 14; Fig 5; 68pp; English.

XX The present invention describes antigenic P5-like fimbria subunit

XX peptides (LB1(f) peptides) of P5-like fimbria proteins from various

XX Haemophilus influenzae strains. The peptides are used for diagnosis,

XX prevention, and treatment of Haemophilus influenzae infections, such as

XX otitis media, sinusitis, conjunctivitis, or lower respiratory tract

XX infection. The peptides may also be used in vaccines against H.

XX influenzae. Antibodies and probes from the present invention can be used

XX for diagnosis of H. influenzae infection. AAY79955 to AAY79993, and

XX AAY91201 to AAZ91252, represent sequences used in the exemplification of

```
CC the present invention
XX
SQ Sequence 464 AA;

Query Match      100.0%; Score 2477; DB 3; Length 464;
Best Local Similarity 100.0%; Pred. No. 5.1e-207;
Matches 464; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
QY 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
DB 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
QY 121 MTENFETKDGKQAQVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIPEIKAPW 180
DB 121 MTENFETKDGKQAQVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIPEIKAPW 180
QY 161 FHHQNGKDIAAETLKVLYKGYDKTDVYLTQDFNFKRIKTELLPQMGMDLKLVLQI 240
DB 161 FHHQNGKDIAAETLKVLYKGYDKTDVYLTQDFNFKRIKTELLPQMGMDLKLVLQI 240
QY 241 AYTDWKETOEDPKGYWVNNYNDWMPKPGAMAEVVKYADGVPGWYMLVNKEESKPDNIV 300
DB 241 AYTDWKETOEDPKGYWVNNYNDWMPKPGAMAEVVKYADGVPGWYMLVNKEESKPDNIV 300
QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVTFDPDTGVEFL 360
DB 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVTFDPDTGVEFL 360
QY 361 KGKSMGCGKAGVALVRSDYKLYNKNSSNSTLKNLGEHRRARAMDGGKAGVALVRSDYK 420
DB 361 KGKSMGCGKAGVALVRSDYKLYNKNSSNSTLKNLGEHRRARAMDGGKAGVALVRSDYK 420
QY 421 FYEDANGTRDHKKGRHTARTSRSDYKFDNKRIDSTSGHHHHH 464
DB 421 FYEDANGTRDHKKGRHTARTSRSDYKFDNKRIDSTSGHHHHH 464

RESULT 2
AAR20108
ID AAR20108 standard; protein; 364 AA.
XX
AC AAR20108;
DT 01-APR-1992 (first entry)
DE Protein D - human Igd-receptor.
XX
KW Immunoglobulin; binding protein; probe; primer; detection; vaccine.
XX
OS Haemophilus influenzae.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Peptide /label= sig_peptide
FT Peptide 16..19
FT /note= "bacterial lipoprotein consensus sequence"
XX
PN W09118926-A.
XX
PD 12-DEC-1991.
XX
PF 31-MAY-1990; 90SE-00001949.
XX
PR 31-MAY-1990; 90SE-00001949.
XX
PA (FORS/) FORSGREN A.
XX
PI Forsgren A;
XX
```

```
DR WPI; 1992-007432/01.
XX N-ESDB; RAQ20265.
XX Surface-exposed protein conserved in strains of Haemophilus influenzae -
PT useful in vaccines, and antibodies to protein, are used to detect
PT Haemophilus in samples.
XX
PS Disclosure; Fig 9a-b; 42pp; English.
XX
CC The surface-exposed protein, conserved in many strains of H. influenzae
CC or related Haemophilus species, has a mol. wt. of 42 kD and a binding
CC capacity for human Igd. The DNA sequence can be used to design probes and
CC primers for the detection of H.influenza and related species. The protein
CC can be used to detect and isolate Igd and to prepare vaccines against
CC H.influenzae
XX
SQ Sequence 364 AA;

Query Match      77.5%; Score 1920; DB 2; Length 364;
Best Local Similarity 99.5%; Pred. No. 1.4e-158;
Matches 362; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
QY 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
DB 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
QY 121 MTENFETKDGKQAQVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIPEIKAPW 180
DB 121 MTENFETKDGKQAQVYNNRFPPLWKSHFRHTFDEIEFIOGLEKSTGKVGVIPEIKAPW 180
QY 181 FHHQNGKDIAAETLKVLYKGYDKTDVYLTQDFNFKRIKTELLPQMGMDLKLVLQI 240
DB 181 FHHQNGKDIAAETLKVLYKGYDKTDVYLTQDFNFKRIKTELLPQMGMDLKLVLQI 240
QY 241 AYTDWKETOEDPKGYWVNNYNDWMPKPGAMAEVVKYADGVPGWYMLVNKEESKPDNIV 300
DB 241 AYTDWKETOEDPKGYWVNNYNDWMPKPGAMAEVVKYADGVPGWYMLVNKEESKPDNIV 300
QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVTFDPDTGVEFL 360
DB 301 YTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNMQYDALLNKSAGTGVTFDPDTGVEFL 360
QY 361 KGIX 364
DB 361 KGIX 364

RESULT 3
ABB84614
ID ABB84614 standard; protein; 358 AA.
XX
AC ABB84614;
XX
DT 24-FEB-2003 (first entry)
DE
DE E. coli glycerophosphoryl phosphodiester phosphodiesterase.
XX
KW NARC10; NARC16; cytostatic; immunosuppressive; dermatological; cardiac;
KW antiinflammatory; nephrotropic; anti-HIV; neuroprotective;
KW antianaemic; cerebroprotective; vasotropic; antidiabetic; anticonvulsant;
KW immunosuppressive; thyromimetic; immunostimulant; antinemic; hypotensive;
KW tranquiliser; neuroleptic; gene therapy; gene mapping; apoptosis;
KW AIDS; cell cycle disruption; programmed cell death regulation;
KW viral infection; nucleosome assembly; phosphate homeostasis;
KW cell cycle regulation; cancer; follicular lymphoma; carcinoma;
KW p53 mutation; graft rejection; hormone-dependent tumour;
KW autoimmune disorder; valvular heart disease;
KW systemic lupus erythematosus; diabetes; Hashimoto's thyroiditis;
KW immune-mediated glomerulonephritis; virus-induced lymphocyte depletion;
```

acquired immunodeficiency syndrome; neurodegenerative disease; stroke; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; spinal muscular atrophy; retinitis pigmentosa; cerebellar degeneration; myelodysplastic syndrome; ischaemic injury; myocardial infarction; reperfusion injury; liver disease; idiopathic dilated cardiomyopathy; ischaemic cardiomyopathy; aplastic anaemia; chronic neutropenia; mania; myelodysplastic syndrome; central nervous system disorder; anxiety; senile dementia; Huntington's disease; hypertension; schizophrenia; severe bipolar affective disorder; glycerophosphoryl phosphodiester phosphodiesterase.

Escherichia coli.

WO200281516-A2.

17-OCT-2002.

16-JAN-2002; 2002WO-US001098.

16-JAN-2001; 2001US-0262306P.

15-JAN-2002; 2002US-00047855.

(MILL-) MILLENNIUM PHARM INC.

Chiang LW;

WPI; 2003-058503/05.

Novel isolated programmed cell death-related polypeptide, NARC10 and NARC16, useful for treating disorders associated with abnormal apoptotic process e.g. Alzheimer's disease, cancer, myocardial infarction, stroke.

Disclosure; Fig 3; 123pp; English.

This invention describes novel cell death-related polypeptides NARC10 and NARC16, located on chromosome 4q11-4q21 and which have cardiant, antiHIV, CC immunosuppressive, dermatological, antiinflammatory, cerebroprotective, CC nootropic, neuroprotective, antianaemic, cardiant, vasotropic, antianemic, CC antidiabetic, immunosuppressive, cytostatic; thyromimetic, nephrotropic, CC immunostimulant, anticonvulsant, tranquiliser, hypotensive and CC neuroleptic activity and can be used in gene therapy. The products of the invention can be used to modulate NARC10 or NARC16 polypeptides or CC polynucleotides, to map NARC genes on a chromosome, e.g. to locate gene CC regions associated with genetic disease or to associate NARC10 or NARC16 CC with a disease. The polypeptides are also useful for modulating the CC apoptotic process, and are therefore useful for modulating, and treating CC disorders associated with increased apoptosis, inhibition of apoptosis or CC disruptions in cell cycle, for regulating cellular functions including CC programmed cell death, nucleosome assembly, phosphate homeostasis and the CC cell cycle. Preferably, the products of the invention are useful for CC treating disorders associated with abnormally low rate or abnormally high CC rate of apoptosis e.g. cancers including follicular lymphomas, carcinomas CC with p53 mutations, or hormone-dependent tumours, autoimmune disorders CC including systemic lupus erythematosus, diabetes, graft rejection, CC Hashimoto's thyroiditis and immune-mediated glomerulonephritis and viral CC infections e.g. infections caused by herpes viruses, virus-induced CC lymphocyte depletion (including acquired immunodeficiency syndrome CC (AIDS)), neurodegenerative diseases manifested by loss of specific sets CC of neurons (including Alzheimer's disease, Parkinson's disease, CC amyotrophic lateral sclerosis, spinal muscular atrophy, retinitis CC pigmentosa, and cerebellar degeneration), myelodysplastic syndromes CC (including aplastic anaemia), ischaemic injuries (including myocardial CC infarction, stroke and reperfusion injury), and toxin (e.g. alcohol) CC induced liver disease, idiopathic dilated cardiomyopathy, ischaemic CC cardiomyopathy and valvular heart disease, aplastic anaemia, chronic CC neutropenia, and myelodysplastic syndromes, central nervous system CC disorders, senile dementia, Huntington's disease, hypertension, CC schizophrenia, attention deficit disorder, mania, anxiety, severe bipolar CC affective disorder (BP-I). This sequence represents the E. coli CC glycerophosphoryl phosphodiester phosphodiesterase protein described in CC the method of the invention

Sequence 358 AA;

Query Match 50.5%; Score 1251.5; DB 6; Length 358;
Best Local Similarity 64.1%; Pred. No. 2.7e-100;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;
Qy 4 KTLALSLLAGVLACGSSHSNMANTQKSDKIIIAHRGASGYLPEHTLESKALAFACQA 63
Db 6 KNLNMAIMMSTIVMGSSAMAD-----SNEKIVIAHRGASGYLPEHTLPKAWAYACQA 59
Qy 64 DYLEQDLAMTKDGRVVIHDFLDGLTDVAKKPPHRRKDCGRYYVIDFTLKEIOSLEMT 123
Db 60 DYLEQDLVMTKDDNLVLDHLYLDRVTDVADRRFDRARKDCGRYYVIDFTLKEISKLFTE 119
Qy 124 NFETKDGQAQVYENRPLMKSHRIHTFDEIEIFIQGLEKSTCKKVIYIPEIKAPWPHH 183
Db 120 GFDIENGKKVQTYPRFPFMGKSDPRVHTFEEIEIFVQGLNHSSTGKNIYIPEIKAPWPHH 179
Qy 184 QNGKIDIAETLVLKKGDKKTDWVYLQTPDFNKLKRIKTELLPQMGMDLVLQLIAYT 243
Db 180 QEGKIDIAETLVLKKGDKKTDWVYLQTPDFNKLKRIKTELLPQMGMDLVLQLIAYT 239
Qy 244 DWKETQSKDPKGYWVNTYNDMMFKPGAMAEVYKYADGVGPGWYMLVNBKESKPDNIWVTP 303
Db 240 DWNETQSKDPKGYWVNTYNDMMFKPGAMAEVYKYADGVGPGWYMLVNBKESKPDNIWVTP 298
Qy 304 LVKELAQYNVEVHYTVRKDALPEFFTDVNMQYDALLNKSGATGVTDFDTGVEFL 360
Db 299 MVQDAQNKLVHPYTVRSKPLPEYTPDVNLQYLNKAGVNGLFTDFDPKAVKFL 355
RESULT 4
ABO61848
ID ABO61848 standard; protein; 359 AA.
XX AC ABO61848;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polypeptide seqid 8365..
XX KW Recombinant expression vector; transcription regulatory element;
XX KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX OS Klebsiella pneumoniae.
XX FN US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Breton GL, Osborne M;
XX WPI; 2003-895346/82.
XX N-PSDB; ACH95399.
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 8365; 932pp; English.
XX The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This is the amino acid sequence of a Klebsiella pneumoniae polypeptide of the invention

SQ Sequence 359 AA;
Query Match 50.5%; Score 1251.5; DB 7; Length 359;
Best Local Similarity 65.3%; Pred. No. 2.7e-100;
Matches 231; Conservative 48; Mismatches 74; Indels 1; Gaps 1;
QY 7 ALSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYL 66
DB 3 AMWMLTALMSGMILSSSALCFESATAADKRWIAHRGASGYLPEHTLPKAWAYAQGADYL 62
QY 67 EQDLAMTKDGRVLVHIDHFLDGLTDVAKKFPFHRHRCGRYYVIDFTLKEIOSLEMTENFE 126
PB 63 EQDLVMTKDDRLVLDHVLDRVTDVAQRFPORARKDGRFYAIDFTLDEIKSLKFTGCFE 122
QY 127 TYGDKQAQVYNNRPFPLWKSHPRIHTFEDEIFBIOGLEKSTGKKVGIYPIEIKAPWPHQNG 186
DB 123 PKNGKVVQYTPGRFPMGKSDPRIHTFEDEIFBIOGLEKSTGKNTGIYPIEIKAPWPHQNG 182
QY 187 KDIAAETLKVLYKKYGYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWK 246
DB 183 KDIAASTLKVLYKEYGYTSKQVYLCDFDANELKRIKNELEPKMGMDLNLVLQIAYTDWN 242
QY 247 ETQEKDPKGYWVNNYNDWMFKPGAMAEVVKYADGVGPGWMLVNKESKPDNIIVTPLVK 306
DB 243 ETQOKQADGKWNYSYNDWMFKPGAMAQIAQYADGIGDPYHMLV-AEGSKPGAVKLTAMVK 301
QY 307 ELAQYNVEHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDTGVFEFL 360
DB 302 EAHASHLQVHPYTVRADQLPEYATNVNQLYDLYNQAGVGDGLFTDFPDKAVQFL 355

RESULT 5
ADF05718
ID ADF05718 standard; protein; 364 AA.
AC ADF05718;
XX-
DT 12-FEB-2004 (first entry)
XX
DE Bacterial polypeptide #1831.
XX
KW Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant.
XX
OS Proteus mirabilis.
XX
PN US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2003-895291/82.
DR N-PSDB; ADF01546.
XX
PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
PS Disclosure; SEQ ID NO 6003; 870pp; English.
XX
CC The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis

CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
SQ Sequence 364 AA;

Query Match 49.2%; Score 1219.5; DB 7; Length 364;
Best Local Similarity 64.7%; Pred. No. 1.7e-97;
Matches 229; Conservative 50; Mismatches 70; Indels 5; Gaps 3;
QY 8 LSLLAAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLE 67
DB 12 LKPLVAGVLLTLS--LSTIA--QAASDKVVIHRGASGYLPEHTLPKALAYAQGADYLE 67
QY 68 QDLAMTKDGRVLVHIDHFLDGLTDVAKKFPFHRHRCGRYYVIDFTLKEIOSLEMTENPET 127
DB 68 QDLVMTKDNQNLVLDHVLDRVTDVAERYPVRAKDKGRYYAIDFTLDEIKLKEFTGFDI 127
QY 128 KQKQAQVYNNRPFPLWKSHPRIHTFEDEIFBIOGLEKSTGKKVGIYPIEIKAPWPHQNGK 187
DB 128 VDGKVKQSYNNRPFPMGKSDPRIHTFQEBIFBIOGLNKSTGQDIGIYPIEIKAPWPHQEGK 187
QY 188 DIAAETLKVLYKKYGYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKE 247
DB 188 DITKVLVLEVLKQYGYTQKSDNVYLQSPDNDLKRIKTELLPQMGMDLKLVLQIAYTDWNE 247
QY 248 TOEKDPKGYWVNNYNDWMFKPGAMAEVVKYADGVGPGWMLVNKESKPDNIIVTPLVKE 307
DB 248 TYEKQPDGTTWNSYNDWMFKPGAMKEIATYADGIGDPYHMLV-EEDSTPDKITLTGMAAD 306
QY 308 LAQYNVEHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDTGVFEFLK 361
DB 307 ARANKLTHTPVTVRVDKLPKYAKDGLYDIYNQAGAGVFTDPDPLGVKFLQ 360

RESULT 6
ABM68758
ID ABM68758 standard; protein; 359 AA.
XX
AC ABM68758;
XX
DT 20-NOV-2003 (first entry)
XX
DE Photorhabdus luminescens protein sequence #1855.
XX
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.
XX
OS Photorhabdus luminescens.
XX
PN WO200294867-A2.
XX
PD 28-NOV-2002.
XX
PF 07-FEB-2002; 2002WO-IB003040.
XX
PR 07-FEB-2001; 2001FR-00001659.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX
DR WPI; 2003-148459/14.
XX
PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

ID AAY06589 standard; protein; 451 AA.
 AC AAY06589;
 DT 26-OCT-1999 (first entry)
 DE Lipoprotein D-MAGE-3-His fusion protein.
 * KW MAGE-3; lipoprotein D; LPD-MAGE-3-His; fusion protein; tumour; melanoma;
 KW breast cancer; bladder cancer; lung cancer;
 KW head and squamous cell carcinoma; colon cancer; oesophagus carcinoma;
 KW vaccine; human.
 XX
 OS Haemophilus influenzae.
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 443
 FT /notes="this residue is additional to the residues
 FT deduced from the nucleotide sequence of AAX87588"
 XX
 PN WO9940188-A2.
 XX
 PD 12-AUG-1999.
 XX
 PF 02-FEB-1999; 99WO-EP000660.
 XX
 PR 05-FEB-1998; 98GB-00002543.
 PR 06-FEB-1998; 98GB-00002650.
 XX*
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Cabezón Silva T, Cohen J, Slaoui MM, Vinals Bassols C;
 XX
 DR WPI; 1999-494293/41.
 DR N-PSDB; AAX87588.
 XX
 PT New protein derivatives used in cancer vaccine therapy for treating a
 PT range of cancers including melanomas, carcinomas and cancers of breast.
 XX
 PS Example 1; Page 64-65; 72pp; English.
 XX
 CC This sequence represents a novel fusion protein composed of lipidated
 CC protein D (LPD) of Haemophilus influenzae B, the human MAGE-3 tumour-
 CC associated antigen and a hexahistidine tail. A vector designed for
 CC recombinant expression of the fusion protein is provided. MAGE-3 cDNA was
 CC amplified using primers that altered the first 5 codons to Escherichia
 CC coli codon usage. The LPD moiety provided the fusion protein with
 CC additional exogenous T-cell epitopes and also increased expression levels
 CC in E. coli. The lipid tail ensured optimal presentation of the antigen to
 CC antigen-presenting cells. The affinity tag facilitated purification. The
 CC invention relates to MAGE proteins fused to an immunological fusion
 CC partner, e.g. LPD-MAGE-3-His. These novel fusion proteins provide
 CC vaccines for immunotherapy of melanomas or other MAGE-associated tumours
 CC like breast, bladder, lung and non-small cell lung cancer, head and
 CC squamous cell carcinoma, colon carcinoma and oesophagus carcinoma
 XX
 SQ Sequence 451 AA;
 Query Match 26.7%; Score 661.5; DB 2; Length 451;
 Best Local Similarity 37.5%; Pred. No. 1.1e-48;
 Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;
 XX
 QY 1 MDPKTLALSLLAAGVLAGCSSHSSNNWANTQMSDKIIIAHRGASGLPPEHTLESKALAPA 60
 DB 1 MDPKTLALSLLAAGVLAGCSSHSSNNWANTQMSDKIIIAHRGASGLPPEHTLESKALAPA 60
 QY 61 QOADYLEQDLAMTKDGRLLVVIHDFLDGLTDVAKKFPFPHRRKDGYYVIDFTLKEIQSLE 120
 DB 61 QOADYLEQDLAMTKDGRLLVVIHDFLDGLTDVAKKFPFPHRRKDGYYVIDFTLKEIQSLE 120

121 MTENFETKDGKQAQVYNNRFPFLWKSHPRIRHFEDEIEFIQGLEKSTGKVGIIYPIEKAPW 180
 121 MTENFETMDLEQRSQ-----HCKPEE-----GLB-ARGEALGLV-GAQAPA 159
 181 FHHQNGKIDIAETLKVLLKYGDKKTDVMVYLQTFDFNELKRIKTELLPQMGMDL-KLVOL 239
 160 TEEQEAASSSTLVEV-----TLGEVPAAESPPQSPQGASSLPTT 201
 240 IAYTDWKETQ-----KDPKGVVYNNYDWMFKPG---AMAEVV-----KY-----AD 279
 202 MNVPLMSQSDSDSNQEEGPFDPDSEFQAALSRAKVAELVHFLLLKTRAREPVTYKAE 261
 280 GVGP---GWYMLNVNKKSKPDNIYVTPLVKELAQYNNVEHPYTVRKDALPEFFFDVNNQMY 336
 262 MLGSVGVNMQYFFPPIEFISKASSSLQLVFGIELMEVDPIGHLYI-----FATCLGLSY 313
 337 DALLNKSGATGVFTDPDPTGVFEL-KGIKSMGDKAGVALVRSYDKLY-----NKNSSN 390
 314 DGLIGDNQI-----MPKAGLLIIVLAIAREGCAPEEKIWEELSLEVFEGREDSITLG 367
 391 STLKNLGEHRRARAMDGGKAGVALVRSYDKFYEDANG-----TRDHKK 433
 368 DPKLLTQH-----FVQENYLEYRQVPGSDPACYEFLMGPRALVETSYKV 413
 434 GRHTARTSRSDYKFD-----NKRIDSTSGHHHHH 464
 414 LHMVVKISGGPHISYPPPLHEWVLREGETSGHHHHH 451
 RESULT 9
 ABU04447
 ID ABU04447 standard; protein; 451 AA.
 XX
 AC ABU04447;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1113.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chicx RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1113; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (BPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 451 AA;

Query Match 26.7%; Score 661.5; DB 6; Length 451;
Best Local Similarity 37.5%; Pred. No. 1.1e-48;
Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;
QY 1 MDPKTLALSLAAGVLAGCSSHSSNMANTOMKSKIIIAHRCASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLAAGVLAGCSSHSSNMANTOMKSKIIIAHRCASGYLPEHTLESKALAPA 60
QY 61 QQADYLEODLANTKQRLVVIHDFDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
DB 61 QQADYLEODLANTKQRLVVIHDFDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
QY 121 MTENFETMDLEORSQ-----HCKPEE-----GLE-ARGALGLV-GAQA 159
DB 121 MTENFETMDLEORSQ-----HCKPEE-----GLE-ARGALGLV-GAQA 159
QY 181 FHHQNGKQIAETLKVLYKGYDKTDMVYLOTFDFNELKRIKTELLPOMGMDL-KLVOL 239
DB 181 FHHQNGKQIAETLKVLYKGYDKTDMVYLOTFDFNELKRIKTELLPOMGMDL-KLVOL 239
QY 160 TBEQEAASSSTLVEV-----TLGEVPAESPPDPQSGASSLPTT 201
DB 160 TBEQEAASSSTLVEV-----TLGEVPAESPPDPQSGASSLPTT 201
QY 240 IAYTWKETQE-----KDPKGYVWVNYDMMKPEG---AMAEVV-----AD 279
DB 240 IAYTWKETQE-----KDPKGYVWVNYDMMKPEG---AMAEVV-----AD 279
QY 202 MNYPLWSQSYEDSSNQEBGPTFPDLSEFOALSRKVAELVHFLLLKYRAREPVTXAE 261
DB 202 MNYPLWSQSYEDSSNQEBGPTFPDLSEFOALSRKVAELVHFLLLKYRAREPVTXAE 261
QY 280 GYGP---GWYMLVNKEESPDNIVTPLVKEIAQYNVEVHPYTVRKDALPEFFTDVQMY 336
DB 280 GYGP---GWYMLVNKEESPDNIVTPLVKEIAQYNVEVHPYTVRKDALPEFFTDVQMY 336
QY 262 MLGSVVGWVQYFPFIFSKASSLQVFGIELMEVDPIGHLI-----PATCGLSY 313
DB 262 MLGSVVGWVQYFPFIFSKASSLQVFGIELMEVDPIGHLI-----PATCGLSY 313
QY 337 DALLNKGATGVTFDPDTGVBEFL-KGIKSMDDGGKAGVALVRSDYKLY-----NKNSSN 390
DB 337 DALLNKGATGVTFDPDTGVBEFL-KGIKSMDDGGKAGVALVRSDYKLY-----NKNSSN 390
QY 314 DGLLGDNQI-----MPKAGLLIIVLAIITAREGDCAPBEKIWEELSVLEFVEGREDSILG 367
DB 314 DGLLGDNQI-----MPKAGLLIIVLAIITAREGDCAPBEKIWEELSVLEFVEGREDSILG 367
QY 391 STLKNLGEHHRARAMDGGKAGVALVRSDYKLYFEDANG-----TRDHKK 433
DB 391 STLKNLGEHHRARAMDGGKAGVALVRSDYKLYFEDANG-----TRDHKK 433
QY 368 DPKLLLTQH-----FVOENYLEYRQVFGSDPACYBFLMGPRALVETSYKV 413
DB 368 DPKLLLTQH-----FVOENYLEYRQVFGSDPACYBFLMGPRALVETSYKV 413
QY 434 GRHTARTSRDYKFD-----NKRIDSTSGHHHHH 464
DB 434 GRHTARTSRDYKFD-----NKRIDSTSGHHHHH 464
QY 414 LHHWVKISGPHISVPLHEWLVREGEETSGHHHHH 451
DB 414 LHHWVKISGPHISVPLHEWLVREGEETSGHHHHH 451

RESULT 10
AAY02348
ID AAY02348 standard; protein; 144 AA.
XX
AC AAY02348;
XX
DT 17-OCT-2003 (revised)
DT 09-JUL-1999 (first entry)
XX
DE Protein D, the fusion partner for HIV-1 Nef/Tat fusion proteins.
XX
KW HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

KW vaccine; HIV infection; protein D.
XX
OS Haemophilus influenzae; B.
XX
PN WO9916884-A1.
XX
PD 08-APR-1999.
XX
PF 17-SEP-1998; 98WO-BP006040.
XX
PR 26-SEP-1997; 97GB-00020585.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck C, Godart SAG, Marchand M;
XX
XX WPI; 1999-302282/25.
DR N-PSDB; AAX35684.
XX
PT HIV Tat or Nef protein linked to a fusion partner.
XX
PS Example 2; Fig 1; 66pp; English.
XX
CC The present sequence represents protein D. This protein acts as the
CC fusion partner for the fusion proteins of th invention, which also
CC comprise HIV-1 Nef or Tat proteins (or derivative). The fusion protein
CC can be used in a vaccine to prevent HIV infection. (Updated on 17-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 144 AA;
Query Match 26.3%; Score 651; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.7e-48;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPKTLALSLAAGVLAGCSSHSSNMANTOMKSKIIIAHRCASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLAAGVLAGCSSHSSNMANTOMKSKIIIAHRCASGYLPEHTLESKALAPA 60
QY 61 QQADYLEODLANTKQRLVVIHDFDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
DB 61 QQADYLEODLANTKQRLVVIHDFDLGLTDVAKKFPFHRKDGRIYVDFTLKEIQSLE 120
QY 121 MTENFET 127
DB 121 MTENFET 127
RESULT 11
AAY06590
ID AAY06590 standard; protein; 446 AA.
XX
AC AAY06590;
XX
DT 26-OCT-1999 (first entry)
XX
DE Lipoprotein D-MAGE-1-His fusion protein.
XX
KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour; melanoma;
KW breast cancer; bladder cancer; lung cancer;
KW head and squamous cell carcinoma; colon cancer; oesophagus carcinoma;
KW vaccine; human.
XX
OS Haemophilus influenzae.
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
PN WO9940188-A2.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-EP000660.

```

XX 05-FEB-1998; 98GB-00002543.
PR 06-FEB-1998; 98GB-00002650.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX Cabezón Silva T, Cohen J, Slaoui WM, Vinals Bassols C;
XX WPI; 1999-494293/41.
DR N-PSDB; AAX87591.
XX
XX New protein derivatives used in cancer vaccine therapy for treating a
PT range of cancers including melanomas, carcinomas and cancers of breast.
XX Example 6; Page 67-68; 72pp; English.
XX
XX The present sequence represents a novel fusion protein composed of
CC lipidated protein D (LPD) of Haemophilus influenzae B, the human MAGE-1
CC tumour-associated antigen and a hexahistidine tail. The invention relates
CC to MAGE proteins fused to an immunological fusion partner such as LPD.
CC The LPD moiety provides the fusion protein with additional exogenous T-
CC cell epitopes and also increase expression levels in E. coli. The lipid
CC tail ensures optimal presentation of the antigen to antigen-presenting
CC cells. The affinity tag facilitates purification. The novel fusion
CC proteins provide vaccines for immunotherapy of melanomas or other MAGE-
CC associated tumours like breast, bladder, lung and non-small cell lung
CC cancer, head and squamous cell carcinoma, colon carcinoma and oesophagus
CC carcinoma
XX
XX Sequence 446 AA;
SQ
Query Match 26.3%; Score 651; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 8.8e-48;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPKTLALSLLAAGVLACGSSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
DB 1 MDPKTLALSLLAAGVLACGSSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
QY 61 QOADYLEQDLAMTKDGRLLVWIDHFDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLE 120
DB 61 QOADYLEQDLAMTKDGRLLVWIDHFDLGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLE 120
QY 121 MTENPET 127
DB 121 MTENPET 127
RESULT 12
AAY02352
ID AAY02352 standard; protein; 324 AA.
XX
AC AAY02352;
XX
DT 09-JUL-1999 (first entry)
XX
DE A representative Lipod-Nef-His fusion protein.
XX
XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;
XX vaccine; HIV infection; protein D.
XX
XX Synthetic.
XX Human immunodeficiency virus 1.
XX
XX WO9916884-A1.
XX
XX 08-APR-1999.
XX
XX 17-SEP-1998; 98WO-EP006040.
XX
XX 26-SEP-1997; 97GB-00020585.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

```

```

XX Bruck C, Godart SAG, Marchand M;
XX WPI; 1999-302282/25.
DR N-PSDB; AAX35688.
XX
XX HIV Tat or Nef protein linked to a fusion partner.
XX
XX Disclosure; Fig 2; 66pp; English.
XX
XX The present sequence represents a fusion protein comprising Lipod-HIV-1
CC Nef-His. The protein is exemplified by the fusion proteins of the invention.
CC The specification also describes fusion proteins comprising HIV-1 Tat
CC protein. The fusion protein can be used in a vaccine to prevent HIV
CC infection
XX
XX Sequence 324 AA;
SQ
Query Match 24.5%; Score 606; DB 2; Length 324;
Best Local Similarity 37.3%; Pred. No. 4.7e-44;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;
QY 19 CSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPAQOQADYLEQDLAMTKDGRLL 78
DB 1 CSHSSNNANTQMSDKIIIAHRGASGYLPEHTLESKALAPAQOQADYLEQDLAMTKDGRLL 60
QY 79 VTIHDFLDGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLENTENFETKDGKQAOVYPN 138
DB 61 VTIHDFLDGLTDVAKKPPHRRKDGRIYVIDFTLKEIQSLENTENFETKDGKQAOVYPN 113
QY 139 RFPLKSHFRHTPEDEIEFTQGLEKSTGKKGVIPIKAPWFHONGKDIAETLKLK 198
DB 114 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
QY 199 KYGVDKTDMVYLTQDFNELKRIKTELLPQMGMDLKLVIATDKETQKDPKGVW 258
DB 148 KHGAITSSNTAATN-----AACAWLEAQEEVEEVPV 179
QY 259 NYNYDNWFKCMAEYVYADGVGPGWMLVNKESKPDNIVTYPVLKELAQYVNEVHPY 318
DB 180 T-----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLHSQR-----214
QY 319 TVRKDALPEFPTDVNQMYDALLNKGATGVFTDF---PDTGVFLKGIKSMGKGAGVA 374
DB 215 --RQDILDLYHTQ-----GYFPDQNYTPGQVRY---PLTFMCYKLV 255
QY 375 LVRSYKLYNNKSSNSTLKNLGHRRARADWGKAGVALVRSYKPYEDANGTRDHKG 434
DB 256 PVEPD-KVEEANKGENTSLHPVSLH---GMDDPPEVLEWRFDRL-----AF 300
QY 435 RHTARTSRSDYKFDYDNKRIDSTSGHHHHH 464
DB 301 HHVARELHPEY--FKN---CTSGHHHHH 324
RESULT 13
AAG63234
ID AAG63234 standard; protein; 324 AA.
XX
XX AAG63234;
XX
DT 01-OCT-2001 (first entry)
XX
DE Amino acid sequence of a His tagged LipD-Nef of HIV.
XX
XX HIV; nef gene; tat gene; vaccine; gp120 gene; HIV viral load.
XX
XX Synthetic.
XX Human immunodeficiency virus.
XX
XX Key Location/Qualifiers
XX Peptide 1..109
XX /note= "ProtD fusion partner"

```


Qy 79 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFITLKEIQSLEMTENFETKDGKQAOQVYPN 138
Db |||||
Qy 61 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFITLKEIQSLEMTENFETKDGKQAOQVYPN 113
Db |||||
Qy 139 RPPLWKSHFRHTFEDEIEFIOGLEKSTGKVGIIPEIKAPWFHONGKIDAAETLKVLE 198
Db |||||
Qy 114 ---W-----SKSSVVG-WPTVRERMRAEPAADGVGAASRDLE 147
Qy 199 KYGVDKKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEDKPKGYVW 258
Db |||||
Qy 148 KHGAITSNTAATN-----AACAWLEAQEEEEVEGPPV 179
Qy 259 NNYNDMMFKPGAMAEVVKYADGVGPMVLYNKEESKPDNIYVTPLVKELAQYNVEHPY 318
Db |||||
Qy 180 T-----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLHSOR----- 214
Qy 319 TVRKDALPEFFTDVNMQYDALLNKGSGATGVFTDF-----PDTGVFELKGIKSMGDKAGVA 374
Db |||||
Qy 215 ---RQDILDLDWIYHTQ-----GYFPDMQNYTPGPGVRY---PLTFGCYKLV 255
Qy 375 LVRSYKLYNKNSSNSTLKNLGEHHRARANDGGKAGVALVRSYKFYEDANGTRDHKKG 434
Db |||||
Qy 256 VPEPD-KYEEANKGENTSLHHPVSLH---GMDDPEREVLWRFDSRL-----AF 300
Qy 435 RHTARTSRDYKFDYDNKRIDSTSGHHHHH 464
Db |||||
Qy 301 HVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 15

AA02354
ID- AA02354 standard; protein; 326 AA.

XX AA02354;

DT 09-JUL-1999 (first entry)

DE A representative Lipod-Nef fusion protein.

XX HIV nef gene; fusion protein; HIV nef gene; Nef protein; Tat protein;

KW vaccine; HIV infection; protein D.

XX Synthetic.

OS Human immunodeficiency virus 1.

XX WO9916884-A1.

PN 08-APR-1999.

PD 17-SEP-1998; 98WO-EP006040.

PF 26-SEP-1997; 97GB-00020585.

PR (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PA Bruck C, Godart SAG, Marchand M;

PI WPI; 1999-302282/25.

XX N-PSDB; AAX35690.

DR HIV Tat or Nef protein linked to a fusion partner.

PT Disclosure; Fig 2; 66pp; English.

PS The present sequence represents a fusion protein comprising Lipod-HIV-1
XX Nef. The protein is exemplified by the fusion proteins of the invention. The
CC specification also describes fusion proteins comprising HIV-1 Tat
CC protein. The fusion protein can be used in a vaccine to prevent HIV
CC infection

XX Sequence 326 AA;

Query Match 24.1%; Score 597; DB 2; Length 326;

Best Local Similarity 37.2%; Pred. No. 2.9e-43;
Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;
Qy 20 SSSNNNANTOMKSDKIIIAHARGASGYLPEHTLESKALAAQADYLEQDLAMTKDGRLV 79
Db 4 SSSNNNANTOMKSDKIIIAHARGASGYLPEHTLESKALAAQADYLEQDLAMTKDGRLV 63
Qy 80 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFITLKEIQSLEMTENFETKDGKQAOQVYPN 139
Db 64 VVIHDFLDGLTDVAKKPPHRRKDGRIYVDFITLKEIQSLEMTENFETKDGKQAOQVYPN 115
Qy 140 FPLWKSHFRHTFEDEIEFIOGLEKSTGKVGIIPEIKAPWFHONGKIDAAETLKVLE 199
Db 116 ---W-----SKSSVVG-WPTVRERMRAEPAADGVGAASRDLE 150
Qy 200 YGYDKKTDVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETOEDKPKGYVW 259
Db 151 HGAISSNTAATN-----AACAWLEAQEEEEVEGPPV 182
Qy 260 YNYDMFKPGAMAEVVKYADGVGPMVLYNKEESKPDNIYVTPLVKELAQYNVEHPY 319
Db 183 ---PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLHSOR----- 216
Qy 320 VRKDALPEFFTDVNMQYDALLNKGSGATGVFTDF-----PDTGVFELKGIKSMGDKAGVAL 375
Db 217 -RQDILDLDWIYHTQ-----GYFPDMQNYTPGPGVRY---PLTFGCYKLV 258
Qy 376 VRSYKLYNKNSSNSTLKNLGEHHRARANDGGKAGVALVRSYKFYEDANGTRDHKKGR 435
Db 259 VEPD-KYEEANKGENTSLHHPVSLH---GMDDPEREVLWRFDSRL-----AFH 303
Qy 436 HTARTSRDYKFDYDNKRIDSTSGHHHHH 464
Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

Search completed: June 22, 2005, 11:37:45

Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:32:20 ; Search time 43 Seconds
(without alignments)
805.515 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLALLAAGVLACG.....YKFYDNKRIDSTSGHHHHH 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1251.5	50.5	359	US-09-489-039A-8365	Sequence 8365, Ap
2	1219.5	49.2	364	US-09-543-681A-6003	Sequence 6003, Ap
3	565	22.8	220	US-09-485-885-1	Sequence 1, Appli
4	565	22.8	220	US-09-485-885-8	Sequence 8, Appli
5	563	22.7	278	US-09-485-885-21	Sequence 21, Appli
6	561.5	22.7	227	US-09-485-885-19	Sequence 19, Appli
7	559.5	22.6	273	US-09-485-885-4	Sequence 4, Appli
8	559.5	22.6	371	US-09-485-885-6	Sequence 6, Appli
9	558	22.5	227	US-09-485-885-16	Sequence 16, Appli
10	558	22.5	383	US-09-485-885-23	Sequence 23, Appli
11	381.5	15.4	386	US-09-328-352-5514	Sequence 5514, Ap
12	381.5	15.4	503	US-09-252-991A-27882	Sequence 27882, A
13	269.5	10.9	316	US-08-956-171E-5231	Sequence 5231, Ap
14	269.5	10.9	316	US-08-781-986A-5231	Sequence 5231, Ap
15	265	10.7	353	US-09-134-001C-3115	Sequence 3115, Ap
16	245.5	9.9	383	US-09-252-991A-25535	Sequence 25535, A
17	201	8.1	260	US-09-134-001C-5097	Sequence 5097, Ap
18	182.5	7.4	359	US-08-457-997B-2	Sequence 2, Appli
19	182.5	7.4	359	US-08-467-722A-2	Sequence 2, Appli
20	182.5	7.4	359	US-09-451-184-2	Sequence 2, Appli
21	173	7.0	247	US-09-107-532A-3856	Sequence 3856, Ap
22	169.5	6.8	203	US-09-710-279-1794	Sequence 1794, Ap
23	149.5	6.0	338	US-08-210-394-1	Sequence 1, Appli
24	148.5	6.0	256	US-09-134-001C-4544	Sequence 4544, Ap
25	143.5	5.8	318	US-09-107-532A-6446	Sequence 6446, Ap
26	133	5.4	382	US-09-134-000C-3829	Sequence 3829, Ap
27	130.5	5.3	256	US-09-583-110-5215	Sequence 5215, Ap

28	130.5	5.3	532	4	US-09-107-433-3404	Sequence 3404, Ap
29	129.5	5.2	331	3	US-09-247-155-116	Sequence 116, App
30	128.5	5.2	364	4	US-09-949-016-10274	Sequence 10274, A
31	126.5	5.1	600	4	US-09-134-000C-5694	Sequence 5694, App
32	118	4.8	253	4	US-09-602-787A-236	Sequence 236, App
33	115	4.6	1308	4	US-09-134-000C-6588	Sequence 6588, Ap
34	114.5	4.6	312	3	US-09-134-001C-3449	Sequence 3449, Ap
35	114	4.6	1529	4	US-09-107-433-4771	Sequence 4771, Ap
36	114	4.6	2138	4	US-09-583-110-5274	Sequence 5274, Ap
37	112	4.5	390	4	US-09-252-991A-18471	Sequence 18471, A
38	111	4.5	253	4	US-09-902-540-11572	Sequence 11572, A
39	108	4.4	655	4	US-09-107-433-3033	Sequence 3033, Ap
40	108	4.4	661	4	US-09-583-110-2948	Sequence 2948, Ap
41	108	4.4	1350	4	US-09-784-5548-4	Sequence 4, Appli
42	108	4.4	10182	3	US-09-134-001C-3159	Sequence 3159, Ap
43	107.5	4.3	252	4	US-09-543-681A-6220	Sequence 6220, Ap
44	107.5	4.3	433	4	US-09-724-623-83	Sequence 83, Appli
45	107.5	4.3	554	1	US-07-839-433-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-489-039A-8365
; Sequence 8365, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8365
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8365

Query Match	50.5%	Score 1251.5;	DB 4;	Length 359;
Best Local Similarity	65.3%	Pred. No. 4.2e+108;		
Matches	231;	Conservative 48;	Mismatches 74;	Indels 1;
Gaps	1;			
Qy	7	ALSLAAGVLACGSSHSNNMANTOMKDKIIIAHRGASGYLPEHTLESKALAPAOQADYL	66	
Db	3	AMKMKLTALMSGMTLSSSALCFSAATAADQWIAHRGASGYLPEHTLPKAWAYAQGADYL	62	
Qy	67	BODLAWTKDGRVVIHDFHLDGLTDVAKKPPHRKDGRIYVDFTLKEIOSLEMTENFE	126	
Db	63	EQDLVMTKDLRVVHDHYLDRTVDVQAFQPARAKDGRFYAIDFTLDEISLKFTEGFE	122	
Qy	127	TKGQKQAVYENRPLMKSHFRIHTFEDIEBFIQGLEKSTGKKGVIYPEIKAPWPHQNG	186	
Db	123	PKNGKNVQTYGCRFPMGKSDPRIHTFEEIEFVQGLNHSHTGKNIGIYPEIKAPWPHQNG	182	
Qy	187	KDIAAETLVKLYGKYDKKTDVYVDFPDLFELKRIKTELLPQMGMDLKLVIAYTQWK	246	
Db	183	KDIAAETLVKLYGKYDKKTDVYVDFPDLFELKRIKTELLPQMGMDLKLVIAYTQWK	242	
Qy	247	ETQSKDPKGVVNNYNDWMPKPGMAEVVYKADGVGPGVYMLVKNKEESKPNIVYTVLVK	306	
Db	243	ETQSKDPKGVVNNYNDWMPKPGMAEVVYKADGVGPGVYMLVKNKEESKPNIVYTVLVK	301	
Qy	307	ELAQYVVEHYPTVYRKDALPEFTFDVNNQYDALLNKSATGVTDFPDPTGVEFL	360	
Db	302	EAHSHLQVHYPTVYRKDALPEFTFDVNNQYDALLNKSATGVTDFPDPTGVEFL	355	

RESULT 2

```
US-09-543-681A-6003
; Sequence 6003, Application US/09543681A
; Patent No. 6505709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6003
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6003

Query Match      49.2%; Score 1219.5; DB 4; Length 364;
Best Local Similarity 64.7%; Pred. No. 4.1e-105; Indels 5; Gaps 3;
Matches 229; Conservative 50; Mismatches 70;

Qy 8 LSLAAGYLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLE 67
Db 12 LKPLVAGVLLTSL--LSIIA--QAASDKVIAHRGASGYLPEHTLPKALAYAQADYLE 67
Qy 68 QDLAMTKDGRLLVVIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPET 127
Db 68 QDLVMTKDNQLVHLHDYLDRTVDAERYPVRARKDGRIYVDFTLKEIQSLEMTENPET 127
Qy- 128 KDGQAQVYVNRFLPKSHFRHTFEDIEFTQLEKSTGKKGVIYPIEIKAPWFHONGK 187
Db 128 VDGKKVQSYNRFPMGSKDFRIHTFQEEIEFTQLEKSTGQDGIYPIEIKAPWFHEQGGK 187
Qy 188 DIAAETLKVLYKKYQKDKTDMVYLQTFDFNFKRIKTELLPQMGNDLKLVLIIATYDWE 247
Db 188 DITKKVLEVLKQYGYTQKSDNVYLSQSPDNDLKRIKTELLPQMGNDLKLVLIIATYDWE 247
Qy 248 TOEKDPKGYVNNYNDWMPKFGAMAEVVKYADGVGPGMYMLVNKESKPDNIVTPLVKE 307
Db 248 TYEQPDGTWNTYSYNDWMPKFGAMAEVVKYADGVGPGMYMLVNKESKPDNIVTPLVKE 306
Qy 308 LAQYVNEVHPYVRKDALPEFTDVMYDALLNKGSGATGVTDFPDGVEFLK 361
Db 307 ARANKLTIHPFTVRVDKLPKYAKOGDQLYDIYNQAGAGVFTDFPDGVRFLQ 360

RESULT 3
US-09-485-885-1
; Sequence 1, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-1

Query Match      22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 44; Gaps 2;
Matches 117; Conservative 3; Mismatches 8;

Qy 20 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLV 79
Db 4 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLV 63
Qy 80 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETK----- 128
Db 64 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETMAHGDPTPTLHE 123
Qy 129 -----DGKQAQVYVNRFLPKSHFRHTF 152
Db 124 YMLDLQPETTDLYCYEQQLNDSSEDEIDGPAGQAEFDR-----AHYNIIVTF 170

RESULT 4
US-09-485-885-8
; Sequence 8, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-8

Query Match      22.8%; Score 565; DB 3; Length 220;
Best Local Similarity 68.0%; Pred. No. 1.7e-44; Indels 44; Gaps 2;
Matches 117; Conservative 3; Mismatches 8;

Qy 20 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLV 79
Db 4 SSSSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLV 63
Qy 80 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETK----- 128
Db 64 VIHDFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENPETMAHGDPTPTLHE 123
Qy 129 -----DGKQAQVYVNRFLPKSHFRHTF 152
Db 124 YMLDLQPETTDLYCYEQQLNDSSEDEIDGPAGQAEFDR-----AHYNIIVTF 170

RESULT 5
US-09-485-885-21
; Sequence 21, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
```

```
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-21

Query Match      22.7%; Score 563; DB 3; Length 278;
Best Local Similarity 34.5%; Pred. No. 3.8e-44;
Matches 157; Conservative 29; Mismatches 79; Indels 190; Gaps 13;

Qy 20 SSSNNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRVL 79
Db 4 SSSNNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRVL 63

Qy 80 VIHDFLDGLTDVAKKPPHRRKDGYYVVDFTLKEIOSLEMTENFET--KDGKQAQVYV- 139
Db 64 VIHDFLDGLTDVAKKPPHRRKDGYYVVDFTLKEIOSLEMTENFET-----111

Qy 140 PFLMKSHFRHTFEDEIEPIQGLESTGKKGVIPEIKAPWPHQNGKDIAAETLKLK 199
Db 112 -----MARFEDPTR-----RPLYLP-----DLCTELNTSLQ- 137

Qy 200 GYDKKTMVYLOTDFNELKRIKTELLPQMGMDLKLVLQIAYTWKQTPQKPGYVNW 259
Db 138 ---DIEITCVCKTY-----LELTV-----208

Qy 260 YNYDMKPGMAEVKYVGVGPGMYMLVNKEESKPDNIYVTPVLKLAQYVNEVHPVT 319
Db 156 ---FEAFKD---LFVTRDSI-----PHAACHKIDFYSIRELRHSDSVYGD 200

Qy 320 VRKDALPEFFTDVQMYDALLNKSGATGVFTDPDTGVFEFLKGISMDGKAGVALVRSD 379
Db 201 LEK-----LNTWG-----208

Qy 380 YKLYN-----KNSSNSTLKNLGEHHRARAMDGGKAGVALVRSDYKFYEDANGTRD 430
Db 209 --LYNLLIRCLRCQPLNPAEKLRLHKKRPHNIAG-----243

Qy 431 HKKGR-HTARTSRSDYKFDNKRIDSTSGHHHHH 464
Db 244 HYGQCHSCCNARQERLQRRETQVTSCHHHHHH 278

RESULT 6
US-09-485-885-19
; Sequence 19, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-4

Query Match      22.6%; Score 559.5; DB 3; Length 273;
Best Local Similarity 91.1%; Pred. No. 7.8e-44;
Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

Qy 20 SSSNNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRVL 79
Db 4 SSSNNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRVL 63

Qy 80 VIHDFLDGLTDVAKKPPHRRKDGYYVVDFTLKEIOSLEMTENFETKDGKQ- AQVYV 138
Db 64 VIHDFLDGLTDVAKKPPHRRKDGYYVVDFTLKEIOSLEMTENFETMAMFQDPQPR 123

Qy 139 RPP 141
Db 124 KLP 126

RESULT 8
US-09-485-885-6
; Sequence 6, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezon Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Bernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
```

```
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-6

Query Match      22.5%; Score 558; DB 3; Length 371;
Best Local Similarity 91.1%; Pred. No. 1.2e-43;
Matches 112; Conservative 1; Mismatches 9; Indels 1; Gaps 1;

QY 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

QY 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFETKDGKQ-AQVYPN 138
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFETMAMFPQDQPRR 123

QY 139 RFP 141
Db 124 KLP 126

RESULT 9
US-09-485-885-16
; Sequence 16, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezón Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-16

Query Match      22.5%; Score 558; DB 3; Length 227;
Best Local Similarity 100.0%; Pred. No. 8.1e-44;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

QY 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 127
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 111
```

```
RESULT 10
US-09-485-885-23
; Sequence 23, Application US/09485885
; Patent No. 6342224
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Cabezón Silva, Teresa
; APPLICANT: Delisse, Anne-Marie Eva Fernande
; APPLICANT: Gerard, Catherine Marie Ghislaine
; APPLICANT: Lombardo-Bencheikh, Angela
; TITLE OF INVENTION: Vaccine
; FILE REFERENCE: B45107
; CURRENT APPLICATION NUMBER: US/09/485,885
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: PCT/EP98/05285
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: GB 9717953.5
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-485-885-23

Query Match      22.5%; Score 558; DB 3; Length 383;
Best Local Similarity 100.0%; Pred. No. 1.8e-43;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSNMANTQMSDKIIIAHKGASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63

QY 80 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 127
Db 64 VIHDFLDGLTDVAKKFPHRKDGRIYVVDFTLKEIOSLEMTENFET 111

RESULT 11
US-09-328-352-5514
; Sequence 5514, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5514
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5514

Query Match      15.4%; Score 381.5; DB 4; Length 386;
Best Local Similarity 31.6%; Pred. No. 5e-27;
Matches 125; Conservative 51; Mismatches 145; Indels 75; Gaps 16;

QY 6 LALSLLAGVLGCGSSH---SSNMANTQMSDKI-IIAHKGASGYLPEHTLESKALAPQA 61
Db 13 LCLSLI8---LVGNCDDDKTETTTTPEYQPKLLVVGHRGASALRPEHTLASQKAI 69

QY 62 QADYLEQDLAMTKDGLRVIIHDFLDGLTDVAKKFPHRHRK-----DG----RY 112
Db 70 GADFIEDLVSTKDGVLVARHENEIGGTTNVSTLSQFADRKTKNIDGVDLTGWFT 129

QY 113 LKEIQSLEMT-----NFKTKGKQAOVYVNPPLWKSHFRHTTFEDEFIOGLEK 169
Db 130 LSELQQLKARERIPFRPANTYNDLYP-----VPTLEQIIELAEANYKTKGI 178
```

QY 170 VGIYPEIKAP-WFHONGKDIAAE--TLKVLKKGVDKKTDMVYLQTFDNELKRIKTEL 226
Db 179 IGLYIETKHTPTFKQN---LWEDTLTKLAKYKTRDIAVPYIQSFVQNLKDKREL 235
QY 227 LPQMGMDLKLVLIAAYTDWKTQEKDPKGYWYNNYDMFKPGMAEVVYKADGVGP--G 284
Db 236 --DLHKTLKHAQIIQLYDSKTSRPADFVESGDTKYADLATAGLKDVKAKYANGVGP 293
QY 285 WYMLVNKESKPDNIYVYPLVKELAQNVVHPYVRKD----- 323
Db 294 YLITFNNDGSYKT---STFISDAHTAGLKVHPYFRPENNELPAPLKCSPDKPAERCPT 349
QY 324 -ALPEFFTDVNMQYDALLNKSGATGVTFDPDTGVE 358
Db 350 GALKEF-----EAY-----FKAGVDGVTFDDPALGRE 376

RESULT 12

US-09-252-991A-27882
; Sequence 27882, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27882
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27882

Query Match 15.4%; Score 381; DB 4; Length 503;
Best Local Similarity 30.5%; Pred. No. 8.4e-27;
Matches 119; Conservative 64; Mismatches 143; Indels 64; Gaps 13;

QY 7 ALSLLAAGVLACSSHSSNM-----ANTQM-----KSKIIIAHRGASGYLPEHTL 52
Db 133 ACLMLAAGLASAAHADTGNVQRAIDWAQQAQSVRHPAGHERSPVIAHRGASGYVPEHTL 192
QY 53 ESKALAFQAQADYLEQDLAMTWKGRVIVHDFLGLTDVAK--KPPHHRK---DG--- 104
Db 193 GAYALAVMVGADYVEPDLVMTDRGKLVARHNDNELGLTTDVAQHPEFADRKTKQYDGVEL 252
QY 105 -RYVVIDFTLKEIOSLEMTENFTKDGQAQVYVNPFRPLWKSHFRHTFDEIFIOGLE 163
Db 253 TGWSEDFTLAEKLTIAERIPT-----IRPGNRL-DGTFEIPQLQEIIDLKSLQ 304
QY 164 KSTGKKGVIYPIKAPWFHONGKDIAAETLKLKKGVDKKTDMVYLQTFDNELKRIK 223
Db 305 ISQORTIGLYPEIKHGTFRQLGLAMERPLVNTLHRNGYLGPRAPVFIQSFVNNLKEK 364
QY 224 TELLPQMGMDLKLVLIAAYTDWKTQEKDPKGYWYNNYDMFKPGMAEVVYKADGVGP 283
Db 365 RL-----TGIRLVQL-----YGSQGPYDQAAGGSLTVAEMATAKRLQVARYAYGVP 413
QY 284 GYMLVNKEES---KPDNIYVYPLVKELAQNVVHPYVR-----KDALPE 327
Db 414 DKSYPVIRPDANGNLGP-----TRFVGDAAHAGLKVHPYTFRAENSLPAEFRSADGNPQ 468
QY 328 FTFDVMQYDALLNKSGATGVTFDPDTGVE 357
Db 469 SRGDLAGEIRAVLD-AGIDGLFSDQPDVAV 497

RESULT 13

US-08-956-171E-5231
; Sequence 5231, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5231:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5231:
US-08-956-171E-5231

Query Match 10.9%; Score 269.5; DB 4; Length 316;
Best Local Similarity 27.5%; Pred. No. 1e-16;
Matches 103; Conservative 48; Mismatches 116; Indels 107; Gaps 16;
QY 8 LSLAAGVLACSSHSSNM-----TQMKSCKI-IIAHRGASGYLPEHTLESKALA 58
Db 30 LSVPTAG-----AEQTNQIANKPOAIQWHTNLTNRFTTIAHRGASGYAPEHTFOAYDKS 84
QY 59 FAQ-QADYLEODLAWTKDGRVIVHDFLGLTDVAKKFPHHRKDGYYVIDFTLKEIQ 117
Db 85 HNELKASYIEDLQRTKDGHLVAMHDETIVNRT-----NNGHKVEDYTLDELK 132
QY 118 SLEMTENFTKDGKQAQV-YENRFPPLWKSHFRHTFDEIFIOGLEKSTGKKGVIYPI 176
Db 133 QLDAGSWFNKYPKVARASYKNA-----KVPTLDEILE-----RYGNANYIET 177
QY 177 KAPFHHQNGKDIAAETLKLKKG-----GYDKTDMVYLQTFDNELKRIKTELLPQMG 232
Db 178 KSPDVY-----PGMEEQLLASLKKHLLNNNNKLNKGNHVMIQSFDSLSKKIHRQ----- 226
QY 233 DLKLVOLIAYTDWKTQEKDPKGYWYNNYDMFKPGMAEVVYKADGVGPQWYMLVNKE 292
Db 227 -NKHVPLVKLVDRKGLQQ-----FNDQRLKRSYAIGLGPD----- 262

```

QY 293 ESKPONIVYTPLVKE-----LAQYNVEVHPYTVRKDALPEFTDVNQMYDAL-INKSGATG 347
      |||::|||
Db 263 -----YTDLTEQNTHHKDLGFIHVHYT-----VNEKADMLRLNKYGVDG 302
      |||::|||
QY 348 VTFDPPDGTGVFLK 361
      |||::|||
Db 303 VETNFADKYKEVIK 316
      |||::|||

RESULT 14
US-08-781-986A-5231
; Sequence 5231, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5231:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5231

```

```

Qy 233 DLKLVLQIAYTDWKEQKDPKGWYVWNYNDWMEKPGMAEIVKYADGVGQWYMLVNKE 292
Db 227 -NKHVPLVKLVDDKGELOQ-----FNDQRLKEIRSYAIGLGD----- 262
Qy 293 ESKPDNIVYPLVKE---LAQNVVEVHPYTVRKDALPEFPTDVMQYDAL-LNKSGATG 347
Db 263 -----YTDLFEQNTHLKDLGFIVHPYT-----VNEKADMLRLNKYGVGD 302
Qy 348 VFTDFPDGTGVFEFLK 361
Db 303 VFTNFADKYKEVIK 316

RESULT 15
US-09-134-001C-3115
; Sequence 3115, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3115
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3115

```

Search completed: June 22, 2005, 11:42:18
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

June 22, 2005, 11:40:51 ; Search time 159 Seconds
(without alignments)
1120.544 Million cell updates/sec

Title: US-09-719-379A-81
 Perfect score: 2477
 Sequence: 1 MDPKTLALSLLAAGVLACGS.....YKFYDNKRIDSTSGHHHHH 464

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 sub

```

Database : Published_Applications AA:*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubppaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubppaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubppaa/US10E_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubppaa/US11A_PUBCOMB.pcp.*
20: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pcp.*
21: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pcp.*
22: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	1251.5	50.5	358	US-10-047-855-12	Sequence 12, Appl
2	661.5	26.7	451	US-10-473-127-1113	Sequence 1113, Ap
3	651	26.3	144	US-10-687-060-7	Sequence 7, Appl
4	606	24.5	324	US-10-203-013-15	Sequence 15, Appl
5	606	24.5	324	US-10-687-060-15	Sequence 15, Appl
6	606	24.5	324	US-10-485-048-15	Sequence 15, Appl
7	597	24.1	326	US-10-203-013-19	Sequence 19, Appl
8	597	24.1	326	US-10-687-060-19	Sequence 19, Appl
9	597	24.1	326	US-10-485-048-19	Sequence 19, Appl
10	594.5	24.0	411	US-10-687-060-17	Sequence 17, Appl
11	594.5	24.0	411	US-10-203-013-17	Sequence 17, Appl

12	594.5	24.0	411	17	US-10-485-048-17	Sequence 17, Appl
13	585.5	23.6	413	14	US-10-203-013-21	Sequence 21, Appl
14	585.5	23.6	413	17	US-10-687-060-21	Sequence 21, Appl
15	585.5	23.6	413	17	US-10-485-048-21	Sequence 21, Appl
16	565	22.8	220	13	US-10-000-903-1	Sequence 1, Appl
17	565	22.8	220	13	US-10-000-903-8	Sequence 8, Appl
18	565	22.8	220	17	US-10-899-771-1	Sequence 1, Appl
19	565	22.8	220	17	US-10-899-771-8	Sequence 8, Appl
20	565	22.8	421	14	US-10-296-770-7	Sequence 7, Appl
21	563	22.7	278	13	US-10-000-903-21	Sequence 21, Appl
22	563	22.7	278	17	US-10-899-771-21	Sequence 21, Appl
23	561.5	22.7	227	13	US-10-000-903-19	Sequence 19, Appl
24	561.5	22.7	227	17	US-10-899-771-19	Sequence 19, Appl
25	559.5	22.6	273	13	US-10-000-903-4	Sequence 4, Appl
26	559.5	22.6	273	17	US-10-899-771-4	Sequence 4, Appl
27	559.5	22.6	371	13	US-10-000-903-6	Sequence 6, Appl
28	559.5	22.6	371	17	US-10-899-771-6	Sequence 6, Appl
29	558	22.5	227	13	US-10-000-903-16	Sequence 16, Appl
30	558	22.5	227	17	US-10-899-771-16	Sequence 16, Appl
31	558	22.5	383	13	US-10-000-903-23	Sequence 23, Appl
32	558	22.5	383	17	US-10-899-771-23	Sequence 23, Appl
33	360.5	14.6	434	15	US-10-424-599-282063	Sequence 282063,
34	360	14.0	393	14	US-10-156-761-14312	Sequence 14312, A
35	347.5	14.0	391	16	US-10-425-115-195137	Sequence 195137,
36	343.5	13.9	391	16	US-10-767-701-44879	Sequence 44879, A
37	346.5	13.2	314	15	US-10-282-122A-46840	Sequence 46840, A
38	321	13.0	416	16	US-10-425-115-352984	Sequence 352984,
39	308	12.4	313	15	US-10-282-122A-45439	Sequence 45439, A
40	296	11.9	395	15	US-10-282-122A-69756	Sequence 69756, A
41	271	10.9	54	14	US-10-153-031-16	Sequence 16, Appl
42	269.5	10.9	309	15	US-10-282-122A-70281	Sequence 70281, A
43	269.5	10.9	316	8	US-08-781-986A-5231	Sequence 5231, Ap
44	269.5	10.9	316	15	US-10-329-624-5231	Sequence 5231, Ap
45	265	10.7	349	15	US-10-282-122A-70498	Sequence 70498, A

ALIGNMENTS

```

RESULT 1
US-10-047-855-12
; Sequence 12, Application US/10047855
; Publication No. US20030165863A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Lillian Wei-Ming
; TITLE OF INVENTION: NARC10 and NARC16, Programmed Cell
; FILE REFERENCE: 35800/242056
; CURRENT APPLICATION NUMBER: US/10/047,855
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,306
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 358
; TYPE: prt
; ORGANISM: Escherichia coli K12
; US-10-047-855-12

```

Db 120 GFDIENGKKVQTPGPFPMGKSDFRVHTFEEBIEFVQGLNHTGKNIGIYFKAIPWFHH 179
Qy 184 QNGKDIAAETLVKLVKYGDKKTDWVYLTQTFDFNELKRIKTELLPOMGMDKLVLQIAYT 243
Db 180 QEGKDIAAETLVKLVKYGDKKTDWVYLTQTFDFNELKRIKTELLPOMGMDKLVLQIAYT 239
Qy 244 DKWTEQEDPKGYWVYNNYDMFKGMAEYVVKVADGVGPGWYMLVKNKEESKPNIVVTP 303
Db 240 DWNETQQQPDGGSWVYNNYDMFKGMAEYVVKVADGVGPGWYMLVKNKEESKPNIVVTP 298
Qy 304 LVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKGATGVFTDFPDTGVVEFL 360
Db 299 MVQDAQKNLVVHPYTVRKDALPEPFTDVNQMYDALLNKGATGVFTDFPDTGVVEFL 355

RESULT 2

US-10-473-127-1113
; Sequence 1113, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1113
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1113

Query Match 26.7%; Score 661.5; DB 16; Length 451;
Best Local Similarity 37.5%; Pred. No. 3.3e-45;
Matches 194; Conservative 51; Mismatches 152; Indels 121; Gaps 18;
Qy 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSKIIIAHKGASGYLPEHTLESKALAPA 60
Db 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSKIIIAHKGASGYLPEHTLESKALAPA 60
Qy 61 QOADYLEODLAWTKDGRVLVHDFHFLDGLTDVAKKFPHRHRKDGYYVVIDFTLKEIOSLE 120
Db 61 QOADYLEODLAWTKDGRVLVHDFHFLDGLTDVAKKFPHRHRKDGYYVVIDFTLKEIOSLE 120
Qy 121 MTENFETDLEQSQ-----HCKPPE-----TLGEVPAAESPPDPQSQGASSLPTT 201
Db 121 MTENFETDLEQSQ-----HCKPPE-----TLGEVPAAESPPDPQSQGASSLPTT 201
Qy 181 FHHQNGKDIAAETLVKLVKYGDKKTDWVYLTQTFDFNELKRIKTELLPOMGMDL-KLVQL 239
Db 160 TEEQEAASSTLVEV-----KDPKGYWVYNNYDMFKPG-----AMAEVW-----AD 279
Qy 240 IAYTDWKEQTQ-----KDPKGYWVYNNYDMFKPG-----AMAEVW-----AD 279
Db 202 MNYPLWSQSYEDSSNQEEGEGSTFPDLESEFQAALSRKVAELVHLLKYPAREPVTYKAE 261
Qy 280 GVGP-----GWYMLVKNKEESKPNIVVTPLVKSLAQYNVEVHPYTVRKDALPEFTDVNQMY 336
Db 262 MLGSVVGNWQYFFPFIKSSSSQLVFGIELMEVDPIGHLIYI-----FATCLGLSY 313

Qy 337 DALLNKSGATGVFTDFPDTGVVEFL-KGIKSMGKGAGVALVRSYDKLY-----NKNSSN 390
Db 314 DGLLGDNIQI-----MPKAGLLIIVLAIAREGDCAPEEKIWEBSLVLEVEGREDLSLG 367
Qy 391 STLNKLGHEHRARAMDGGKAGVALVRSYDKFYEDANG-----TRDHKK 433
Db 368 DPKKLLTQH-----PVQENYLEYRQVPGSDPACYBFLMGPRALVETSYVKV 413
Qy 434 GRHTARTSRSDYKPYD-----NKRIDSTSGHHHHH 464
Db 414 LHHMVKISGGPHISYPLPHEWLVREGEETSGHHHHH 451

RESULT 3

US-10-687-060-7
; Sequence 7, Application US/10687060
; Publication No. US20050033022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; APPLICANT: Marc-Hand, Martine
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/BP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Pichia pastoris
US-10-687-060-7

Query Match 26.3%; Score 651; DB 17; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.6e-45;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSKIIIAHKGASGYLPEHTLESKALAPA 60
Db 1 MDPKTLALSLLAAGVLACGSSSHSNMANTQMSKIIIAHKGASGYLPEHTLESKALAPA 60
Qy 61 QOADYLEODLAWTKDGRVLVHDFHFLDGLTDVAKKFPHRHRKDGYYVVIDFTLKEIOSLE 120
Db 61 QOADYLEODLAWTKDGRVLVHDFHFLDGLTDVAKKFPHRHRKDGYYVVIDFTLKEIOSLE 120
Qy 121 MTENFET 127
Db 121 MTENFET 127

RESULT 4

US-10-203-013-15
; Sequence 15, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1e1 Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200.4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336.9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013806.5

```
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-15

Query Match      24.5%; Score 606; DB 14; Length 324;
Best Local Similarity 37.3%; Pred. No. 6.9e-41;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 19 CSSHSSNNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 78
DB 1 CSSHSSNNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 60

QY 79 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYPN 138
DB 61 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYPN 113

QY 139 RPLWKSFRHITFEDEIEFIOGLEKSTGKVGIYPEIKAPWFHONGKIDIAAETLVK 198
DB 114 -----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147

QY 199 KYGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETQKPKGYWV 258
DB 148 KHGAITSNTAATN-----AACAWLEAQEEVEEVGPPV 179

QY 259 NYNDWMPKPGMAEVKADGVGPGWYMLVNKESKPDNIYVTPLVKELAQYNVEVHPY 318
DB 180 T-----PQVPLRPMTYKAADLSHFL---KEKGGLEGLHSQR----- 214

QY 319 TVRKDALPEFFTDVNMQYDALLNKSGATGVTFDF-----PDTGVFLKGIKSMGGKAGVA 374
DB 215 --RQDILDLYHTQ-----GYPPDQNYTPGPGVRY-----PLTFGWCYKLV 255

QY 375 LVRSYKLYNKNSSNSTLKNLGEHRRARAMDGGKAGVALVRSYKFDYEDANGTRDHKKG 434
DB 256 PVSPD-KVEEANKGENTSLHPVSLH---GMDPPEVLEWRFDSRL-----AF 300

QY 435 RHTARTSRDYKFDYDNKRIDSTSGHHHHH 464
DB 301 HHVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 5
US-10-687-060-15
; Sequence 15, Application US/10687060
; Publication No. US20050033022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; APPLICANT: Marc-Hand, Martine
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/EP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Pichia pastoris

US-10-687-060-15
Query Match      24.5%; Score 606; DB 17; Length 324;
Best Local Similarity 37.3%; Pred. No. 6.9e-41;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 19 CSSHSSNNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 78
DB 1 CSSHSSNNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 60

QY 79 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYPN 138
DB 61 VVIHDHFLDGLTDVAKKPPHRRKDGRIYVDFTLKEIQSLEMTENFETKQKQAVYPN 113

QY 139 RPLWKSFRHITFEDEIEFIOGLEKSTGKVGIYPEIKAPWFHONGKIDIAAETLVK 198
DB 114 -----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147

QY 199 KYGDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYTDWKETQKPKGYWV 258
DB 148 KHGAITSNTAATN-----AACAWLEAQEEVEEVGPPV 179

QY 259 NYNDWMPKPGMAEVKADGVGPGWYMLVNKESKPDNIYVTPLVKELAQYNVEVHPY 318
DB 180 T-----PQVPLRPMTYKAADLSHFL---KEKGGLEGLHSQR----- 214

QY 319 TVRKDALPEFFTDVNMQYDALLNKSGATGVTFDF-----PDTGVFLKGIKSMGGKAGVA 374
DB 215 --RQDILDLYHTQ-----GYPPDQNYTPGPGVRY-----PLTFGWCYKLV 255

QY 375 LVRSYKLYNKNSSNSTLKNLGEHRRARAMDGGKAGVALVRSYKFDYEDANGTRDHKKG 434
DB 256 PVSPD-KVEEANKGENTSLHPVSLH---GMDPPEVLEWRFDSRL-----AF 300

QY 435 RHTARTSRDYKFDYDNKRIDSTSGHHHHH 464
DB 301 HHVARELHPEY--FKN-----CTSGHHHHH 324

RESULT 6
US-10-485-048-15
; Sequence 15, Application US/10485048
; Publication No. US20050058657A1
; GENERAL INFORMATION:
; APPLICANT: ERTL, Peter Franz
; APPLICANT: TITE, John Philip
; APPLICANT: VAN WELY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GPI20 and NEF and/or
; TITLE OF INVENTION: TAT for the Immunisation Against HIV
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-048-15

Query Match      24.5%; Score 606; DB 17; Length 324;
Best Local Similarity 37.3%; Pred. No. 6.9e-41;
Matches 168; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 19 CSSHSSNNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 78
DB 1 CSSHSSNNANTQMSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGRLL 60
```

[illegible]

```

RESULT 7
US-10-203-013-19
; Sequence 19, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1el Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200.4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336.9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013806.5
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-19

```

Query Match	24.1%	Score	597	DB	14	Length	326
Best Local Similarity	37.2%	Pred.	No.3.8e-40				
Matches	167	Conservative	34	Mismatches	118	Indels	130
Gaps	15						
Qy	20	SSHSSNNANTQMKSDKIIIAH	RGASGYLPEHTLESKALAF	AAQADYLEQDLAMTKDGR	LV	79	
Db	4	SSHSSNNANTQMKSDKIIIAH	RGASGYLPEHTLESKALAF	AAQADYLEQDLAMTKDGR	LV	63	
Qy	80	VIHDHFDLGLTDVAKKPPHRR	KDGRVYVDFTLKEIQSL	EMTENFETKDGKQAVV	PNR	139	
Db	64	VIHDHFDLGLTDVAKKPPHRR	KDGRVYVDFTLKEIQSL	EMTENFETMGGK		115	
Qy	140	PPLWKSHRIHTFEDETFEIO	GLESKCGKVGVIIPEIKAP	WFFHQNGKDIAETFLK	VKK	199	
Db	116	---W-----	-----SKSSVVG-WPTVR	RRMRRAEPAADGVGAAS	RDLK	150	
Qy	200	YGDYKKTDMVYVLTQTFD	FNELKRIKTELLPQMGMD	KLVLQIAYTDWKETQEK	DPKGYWVN	259	

151	Db	HGAITSNTAATN-----AACAWLEAQEEEEVGPPVT	182
260	Qy	YNIDMFKPGAMAEVVKYADGVGFWYMLNKESKPDNIIVYTVPLKELAQYNVEVHPYT	319
183	Db	-----PQVPLRPMTYKAAVDLSHFU-----KEKGGLEGLIHSQR-----	216
320	Qy	VRKDALPEPFTDNNQMYDALLINKSGATGVFTDF-----PDTGVPEFLKGIKSMDDGGKAGVAL	375
217	Db	-ROILDILWIYHTO-----GYFPDWNQYTFPGPVRY-----PLTFGMWCYKLVLP	258
376	Qy	VRSDYKLYKNSSNSTLKNLGEHRRARAMDDGGKAGVALVRSDYKIFYEDANGTRDHHKGR	435
259	Db	VEPD-KYVEANKGENTSLLHPVSLH---GMDDPEREVLWRFDSRL-----AFH	303
436	Qy	HTARTSRSDYKIFYDNKRIDSTSGHHHHH	464
304	Db	HVARELLHPEY--FKN-----CTSGHHHHH	326

```

RESULT 8
US-10-687-060-19
? Sequence 19, Application US/10687060
? Publication No. US20050033022A1
? GENERAL INFORMATION:
? APPLICANT: Bruck, Claudine
? APPLICANT: Godart, Stephane Andre Georges
? APPLICANT: Marc-Hand, Martine
? TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
? TITLE OF INVENTION: and/or Nef Proteins
? FILE REFERENCE: B45110
? CURRENT APPLICATION NUMBER: US/10/687,060
? CURRENT FILING DATE: 2003-10-16
? PRIOR APPLICATION NUMBER: US/09/509,239
? PRIOR FILING DATE: 2000-03-23
? PRIOR APPLICATION NUMBER: PCT/EP98/06040
? PRIOR FILING DATE: 1998-09-17
? PRIOR APPLICATION NUMBER: GB 9720585.0
? PRIOR FILING DATE: 1997-09-26
? NUMBER OF SEQ ID NOS: 27
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 19
? LENGTH: 326
? TYPE: PRT
? ORGANISM: Pichia pastoris
US-10-687-060-19

```

Query Match	24.1%	Score	597;	DB	17;	Length	326;
Best Local Similarity	37.2%;	Pred. No.	3.8e-40;				
Matches	167;	Conservative	34;	Mismatches	118;	Indels	130;
Gaps	15;						
Qy	20	SSHSSNNANTQMKSDKIIIIAHRGASGYLPBHTTLESKALAFQAQADYLEQDLAMTKDGRLV	79				
Db	4	SSHSSNNANTQMKSDKIIIIAHRGASGYLPBHTTLESKALAFQAQADYLEQDLAMTKDGRLV	63				
Qy	80	VIHDHFLDGLTVAKKFPHRRKDGRIYVYIDFTLKEIQSLEMTENFETGKQKQAVYPNR	139				
Db	64	VIHDHFLDGLTVAKKFPHRRKDGRIYVYIDFTLKEIQSLEMTENFETMGSK-----	115				
Qy	140	FPWLKSHFRITHPEDEIETFTQGLEKSTGKKVGVYPEIKAPWFHHQNGDKIAETLLKVLKK	199				
Db	116	---W-----SKSVVG-WPTVRMRRAEPAADGVGAASRDLEK	150				
Qy	200	YGVDKKTDVYLTQTFDFNELKRIKTELLPQWGMDLKLVQLIAYTDWKETQEKDPKGYWVN	259				
Db	151	HGAITSNTAATN-----AACAWLEAQEEEEVEGPPVT	182				
Qy	260	YNYDMWFKPGAMAEVVKYADGVGPGWYMLVNKBEKSPDNIVYTPVLKELAQYNNVEVHPYT	319				
Db	183	-----PQVPLRPMTYKAAVDLSHFL---KEKGGLEGLIHSQR-----	216				
Qy	320	VRKDALPEFTFDVNMQYDALLNKSGATGVFTDP-----PDTGVFPLKGIKSMDCGKAGVAL	375				
Db	217	-RQDILDLTWYHTQ-----GYPPDMQNYTPGPGVRY-----PLTFPGWCYKLVLP	258				

QY 376 VRSDYKLYNKNSSNSTLKNLGEHHRARAMDGGKAGVALVRSDYKPYEDANGTRDHKKGR 435
Db 259 VEPP-KVEEANKGENTSLHPSVSLH---GMDDPEREVLEWRFSRL-----AFH 303
QY 436 HTARTSRSDYKPYDNKRIDSTSGHHHHH 464
Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

RESULT 9

US-10-485-048-19
; Sequence 19, Application US/10485048
; Publication No. US2005005857A1
; GENERAL INFORMATION:
; APPLICANT: ERTL, Peter Franz
; APPLICANT: TITE, John Philip
; APPLICANT: VAN WELY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-485-048-19

Query Match 24.1%; Score 597; DB 17; Length 326;
Best Local Similarity 37.2%; Pred. No. 3.8e-40;
Matches 167; Conservative 34; Mismatches 118; Indels 130; Gaps 15;

QY 20 SSSSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 79
Db 4 SSSSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLV 63
QY 80 VIHDHFLDGLTDVAKKFPHRHRKGRYVYIDFTLKEIQSLENTENFETKQKQAQVYVNR 139
Db 64 VIHDHFLDGLTDVAKKFPHRHRKGRYVYIDFTLKEIQSLENTENFETMGK----- 115
QY 140 FPLWKSHFRHTFEDEIEFIOGLEKSTGKKGVIYPEIKAPWFHONGKDIAAETLKVLLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
QY 200 YGYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLIAAYTDWKETOEDKPKGYVNV 259
Db 151 HGATSSNTAATN-----AACAWLEAQSEEEVGFVPT 182
QY 260 YNYDMFKPGAMAEVVKYADGVGPWYMLVKNKESSKPDNIYVTPPLVKELAQYVNVVHPT 319
Db 183 -----PQVPLRPMTYKAADVLSHFL---KEKGGLEGLHSQR----- 216
QY 320 VRKDALPEFTDVMQYDALLNKSGATGVFTDF-----PDTGVFLKGIKSMGDKGAGVAL 375
Db 217 -RODILDLMVHTQ-----GYFPDQWNYTPGPGVRY-----PLTFGCYKLV 258
QY 376 VRSDYKLYNKNSSNSTLKNLGEHHRARAMDGGKAGVALVRSDYKPYEDANGTRDHKKGR 435
Db 259 VEPP-KVEEANKGENTSLHPSVSLH---GMDDPEREVLEWRFSRL-----AFH 303
QY 436 HTARTSRSDYKPYDNKRIDSTSGHHHHH 464
Db 304 HVARELHPEY--FKN-----CTSGHHHHH 326

RESULT 10

US-10-203-013-17
; Sequence 17, Application US/10203013
; Publication No. US20030158134A1
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Biologicals S.A.
; TITLE OF INVENTION: No. US20030158134A1e1 Use
; FILE REFERENCE: B45209
; CURRENT APPLICATION NUMBER: US/10/203,013
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: GB 0002200.4
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: GB 0009336.9
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: GB 0013806.5
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: PCT/EP00/05998
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-17

Query Match 24.0%; Score 594.5; DB 14; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
QY 19 CSHSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLR 78
Db 1 CSHSSNMANTOMKSDKIIIAHRCASGYLPEHTLESKALAFQAQADYLEQDLAMTKDGLR 60
QY 79 VIHDHFLDGLTDVAKKFPHRHRKGRYVYIDFTLKEIQSLENTENFETKQKQAQVYVNR 138
Db 61 VIHDHFLDGLTDVAKKFPHRHRKGRYVYIDFTLKEIQSLENTENFETMGK----- 113
QY 139 RFPWLKSHFRHTFEDEIEFIOGLEKSTGKKGVIYPEIKAPWFHONGKDIAAETLKVLLK 198
Db 114 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
QY 199 KYGYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLIAAYTDWK----- 246
Db 148 KHGAITSNTAATNAACAWLEAQSEEEVGFVPTQVPLRPMTYKAAVDLSHFLKEKGGLE 207
QY 247 ---ETQEKDPKGYVW-----NINYDMFKPGAMAEVVKYADGVGP-----GW-YMLVY- 290
Db 208 GLIHSQRQDILDLMVHTQGYFPDQW-----NYTPGPGVRYVPLTFGCYKLV 257
QY 291 -----KEESKPDNIYVTPPLVKELA-----QYNVEVHPVTVRKDALPEPTD 331
Db 258 EPDKVEEANKGENTSLHPSVSLHGMDDPEREVLEWRFSRLAFPHVARELHPEYVKN 314
QY 332 VNQMYDALLNKSGATGVFTDF-----DTGVFLKGIKSMGDKGAGVALVRSDYKLYNKNSSN 390
Db 315 C-----TSEPVDPLRPWKHPCGSP-----KTACTN 340
QY 391 STLKVLGEH-----HPRAMDGGKAGVALVRSDYKPYEDANGTRDHKKGHRTARTSRSD 444
Db 341 CYCKKCCFHCQVCFITKALGISYGR-----KKRRORRRPPQGSQTHQVSLSKOPTSQSR 394
QY 445 YKFYDNKRIDSTSGHHHHH 464
Db 395 ---GDPTGPKETSGHHHHH 411

RESULT 11

US-10-687-060-17
; Sequence 17, Application US/10687060
; Publication No. US2005003022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine

APPLICANT: Godart, Stephane Andre Georges
APPLICANT: Marc-Hand, Martine
TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
TITLE OF INVENTION: and/or Nef Proteins
FILE REFERENCE: B45110
CURRENT APPLICATION NUMBER: US/10/687,060
CURRENT FILING DATE: 2003-10-16
PRIOR APPLICATION NUMBER: US/09/509,239
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/EP98/06040
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: GB 9720585.0
PRIOR FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 411
TYPE: PRT
ORGANISM: Pichia pastoris
US-10-687-060-17

Query Match 24.0%; Score 594.5; DB 17; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
Qy 19 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEODLANTKQRL 78
Db 1 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEODLANTKQRL 60
Qy 79 VVIHDHFDGLDITVAKKPPHRRKDGRIYVIDFTLKEIQSLEMTENFETKDGKQAQVYPN 138
Db 61 VVIHDHFDGLDITVAKKPPHRRKDGRIYVIDFTLKEIQSLEMTENFETMGK----- 113
Qy 139 RPPLWKSHFRHITFEDETEFTQGLEKSTGKVGIYPEIKAPWFHONGKDIATLKVLK 198
Db 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
Qy 199 KGVGDKKTD-----WVYLQTFDENELKRIKTELLPQMGMDLK-LVQLIAYTDWK---- 246
Db 148 KHGAITSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLE 207
Qy 247 ----ETQKDPKGYVW-----NYNDMMFKPGAMAEVVKYADGVGP-----GW-YMLVN- 290
Db 208 GLIHSQRQDILDWIYHTQGYFDMQ-----NYTPGPGVRYPLTFGWCYKLPV 257
Qy 291 -----KEESKPDNIVYPLVKELA-----QYNVEVHPYTVRKDALPEPFTD 331
Db 258 EPDKVVEANKGEN---TSLHPVSLHGMDDPEREVLWRFDRLAFHHVARELHPEYFKN 314
Qy 332 VNQMYDALLNKSGATGVFTDPP-DTGVFELKGIKMDGKAGVALVRSDYKLYNKNSSN 390
Db 315 C-----TSEPVDPRLEPWKPGSQP-----KTACTN 340
Qy 391 STLKNLGEH-----HPRAMDGGKAGVALVRSDYKPYEDANGTRDHHKGRHTARTSRD 444
Db 341 CYCKKCCFHCQVCFTKALGISYGR-----KKRRQRRPPQGSQTHQVSLSKOPTSQSR 394
Qy 445 YKFDVKNKRIDSTSGHHHHH 464
Db 395 ---GDPTGPKETSGHHHHH 411

RESULT 12

US-10-485-048-17
Sequence 17, Application US/10485048
Publication No. US20050058657A1
GENERAL INFORMATION:
APPLICANT: ERTL, Peter Franz
APPLICANT: TITE, John Philip
APPLICANT: VAN WELY, Catherine Anne
APPLICANT: VOSS, Gerald
TITLE OF INVENTION: Vaccine Comprising GP120 and NEF and/or
TITLE OF INVENTION: TAT for the Immunisation Against HIV

FILE REFERENCE: B45279
CURRENT APPLICATION NUMBER: US/10/485,048
CURRENT FILING DATE: 2004-01-26
PRIOR APPLICATION NUMBER: PCT/EP02/08343
PRIOR FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: GB 0118367.2
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 98
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-10-485-048-17

Query Match 24.0%; Score 594.5; DB 17; Length 411;
Best Local Similarity 35.6%; Pred. No. 8.4e-40;
Matches 178; Conservative 45; Mismatches 134; Indels 143; Gaps 19;
Qy 19 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEODLANTKQRL 78
Db 1 CSSHSSNMANTOMKSDKIIIAHRGASGYLPEHTLESKALAFQAQADYLEODLANTKQRL 60
Qy 79 VVIHDHFDGLDITVAKKPPHRRKDGRIYVIDFTLKEIQSLEMTENFETKDGKQAQVYPN 138
Db 61 VVIHDHFDGLDITVAKKPPHRRKDGRIYVIDFTLKEIQSLEMTENFETMGK----- 113
Qy 139 RPPLWKSHFRHITFEDETEFTQGLEKSTGKVGIYPEIKAPWFHONGKDIATLKVLK 198
Db 114 ----W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLE 147
Qy 199 KGVGDKKTD-----WVYLQTFDENELKRIKTELLPQMGMDLK-LVQLIAYTDWK---- 246
Db 148 KHGAITSNTAATNAACAWLEAQEEVEGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLE 207
Qy 247 ----ETQKDPKGYVW-----NYNDMMFKPGAMAEVVKYADGVGP-----GW-YMLVN- 290
Db 208 GLIHSQRQDILDWIYHTQGYFDMQ-----NYTPGPGVRYPLTFGWCYKLPV 257
Qy 291 -----KEESKPDNIVYPLVKELA-----QYNVEVHPYTVRKDALPEPFTD 331
Db 258 EPDKVVEANKGEN---TSLHPVSLHGMDDPEREVLWRFDRLAFHHVARELHPEYFKN 314
Qy 332 VNQMYDALLNKSGATGVFTDPP-DTGVFELKGIKMDGKAGVALVRSDYKLYNKNSSN 390
Db 315 C-----TSEPVDPRLEPWKPGSQP-----KTACTN 340
Qy 391 STLKNLGEH-----HPRAMDGGKAGVALVRSDYKPYEDANGTRDHHKGRHTARTSRD 444
Db 341 CYCKKCCFHCQVCFTKALGISYGR-----KKRRQRRPPQGSQTHQVSLSKOPTSQSR 394
Qy 445 YKFDVKNKRIDSTSGHHHHH 464
Db 395 ---GDPTGPKETSGHHHHH 411

RESULT 13

US-10-203-013-21
Sequence 21, Application US/10203013
Publication No. US20030158134A1
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham Biologicals S.A.
TITLE OF INVENTION: No. US20030158134A1el Use
FILE REFERENCE: B45209
CURRENT APPLICATION NUMBER: US/10/203,013
CURRENT FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: GB 0002200.4
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: GB 0009336.9
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: GB 0013806.5
PRIOR FILING DATE: 2000-06-06
PRIOR APPLICATION NUMBER: PCT/EP00/05998

```
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-203-013-21

Query Match      23.6%; Score 585.5; DB 14; Length 413;
Best Local Similarity 35.5%; Pred. No. 4.6e-39;
Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

QY 20 SSSNNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLEDQDLMATKGRIV 79
Db 4 SSSNNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLEDQDLMATKGRIV 63
QY 80 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETKDGKQAQVYNNR 139
Db 64 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETMGGK----- 115
QY 140 FPLKSHFRIHTFEDEIEFIQGLEKSTGKKGVIYPEIKAPWPHHONGKIDIAAETLKLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
QY 200 GYDVKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246
Db 151 HGAITSSNTAATNAACAWLEAEQEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEG 210
QY 247 ---ETQEKDPKGYV-----NYYDMFKPGMAEVVKYADGVGP-----GW-YMLVN-- 290
Db 211 LIHSQRQDILDLMYHTQGYFPDQ-----NYTPGVRYPPLTFGMCYKLVPE 260
QY 291 ---KEESKPDNIYVTVPLVKELA-----OYNVEVHPYTVRKDALPFEFTDV 332
Db 261 PDKVEANKGEN---TSLHPVSLHGMDDPEREVLEWFRDLSLAFHHVARELHPEYFNK 317
QY 333 NQMDYDALLKSGATGVFTDFF-DTGVEFLKGIKSGMDGKAGVALVRSYKLYNKNSSNS 391
Db 318 -----TSEPVDPRLEPKWHPGSP-----KTACTNC 343
QY 392 TLKNLGEH-----HARAMDGGKAGVALVRSYDYFYEDANGTRDHHKGRHTARTSRDY 445
Db 344 YCKKCCFHCQVCFITKALGISYGR-----KKRRQRRPPQSGSOTHQVSLSKQPTSQSR- 396
QY 446 KFYDNKRIDSTSGHHHHH 464
Db 397 --GDPTGPKETSGHHHHH 413

RESULT 14
US-10-687-060-21
; Sequence 21, Application US/10687060
; Publication No. US2005003022A1
; GENERAL INFORMATION:
; APPLICANT: Bruck, Claudine
; APPLICANT: Godart, Stephane Andre Georges
; APPLICANT: Marc-Hand, Martine
; TITLE OF INVENTION: Fusion Proteins Comprising HIV-1 TAT
; TITLE OF INVENTION: and/or Nef Proteins
; FILE REFERENCE: B45110
; CURRENT APPLICATION NUMBER: US/10/687,060
; PRIOR FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: US/09/509,239
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: PCT/EP98/06040
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: GB 9720585.0
; PRIOR FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 413

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-485-048-21

Query Match      23.6%; Score 585.5; DB 17; Length 413;
Best Local Similarity 35.5%; Pred. No. 4.6e-39;
Matches 177; Conservative 45; Mismatches 134; Indels 143; Gaps 19;

QY 20 SSSNNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLEDQDLMATKGRIV 79
Db 4 SSSNNMANTQMSDKIIIAHAGSGYLPEHTLESKALAFQAQADYLEDQDLMATKGRIV 63
QY 80 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETKDGKQAQVYNNR 139
Db 64 VIHDFLDGLTDVAKFPHRHRKGRYYVIDFTLKEIOSLEMTENFETMGGK----- 115
QY 140 FPLKSHFRIHTFEDEIEFIQGLEKSTGKKGVIYPEIKAPWPHHONGKIDIAAETLKLK 199
Db 116 ---W-----SKSSVVG-WPTVRMRRAEPAADGVGAASRDLEK 150
QY 200 GYDVKTD-----MVYLQTFDFNELKRIKTELLPQMGMDLK-LVQLIAYTDWK----- 246
Db 151 HGAITSSNTAATNAACAWLEAEQEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEG 210
QY 247 ---ETQEKDPKGYV-----NYYDMFKPGMAEVVKYADGVGP-----GW-YMLVN-- 290
Db 211 LIHSQRQDILDLMYHTQGYFPDQ-----NYTPGVRYPPLTFGMCYKLVPE 260
QY 291 ---KEESKPDNIYVTVPLVKELA-----OYNVEVHPYTVRKDALPFEFTDV 332
Db 261 PDKVEANKGEN---TSLHPVSLHGMDDPEREVLEWFRDLSLAFHHVARELHPEYFNK 317
QY 333 NQMDYDALLKSGATGVFTDFF-DTGVEFLKGIKSGMDGKAGVALVRSYKLYNKNSSNS 391
Db 318 -----TSEPVDPRLEPKWHPGSP-----KTACTNC 343
QY 392 TLKNLGEH-----HARAMDGGKAGVALVRSYDYFYEDANGTRDHHKGRHTARTSRDY 445
Db 344 YCKKCCFHCQVCFITKALGISYGR-----KKRRQRRPPQSGSOTHQVSLSKQPTSQSR- 396
QY 446 KFYDNKRIDSTSGHHHHH 464
Db 397 --GDPTGPKETSGHHHHH 413

RESULT 15
US-10-485-048-21
; Sequence 21, Application US/10485048
; Publication No. US20050058657A1
; GENERAL INFORMATION:
; APPLICANT: ERTL, Peter Franz
; APPLICANT: TITE, John Philip
; APPLICANT: VAN WELY, Catherine Anne
; APPLICANT: VOSS, Gerald
; TITLE OF INVENTION: Vaccine Comprising GPI20 and NEF and/or
; TITLE OF INVENTION: Tact for the Immunisation Against HIV
; FILE REFERENCE: B45279
; CURRENT APPLICATION NUMBER: US/10/485,048
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: PCT/EP02/08343
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: GB 0118367.2
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-485-048-21

Query Match      23.6%; Score 585.5; DB 17; Length 413;
Best Local Similarity 35.5%; Pred. No. 4.6e-39;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:31:10 ; Search time 41 Seconds
(without alignments)
1088.892 Million cell updates/sec

Title: US-09-719-379a-81
Perfect score: 2477
Sequence: 1 MDPKTLALSLLAAGVLGCS.....YKFYDNKRIDSTSGHHHHH 464
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1920	77.5	364	2	A43576
2	1916	77.4	364	2	S59932
3	1914	77.3	364	2	S59931
4	1910	77.1	364	2	S59934
5	1902	76.8	364	2	G64086
6	1885	76.1	364	2	S59933
7	1885	76.1	364	2	S59936
8	1261.5	50.9	358	2	D91019
9	1261.5	50.9	358	2	F85863
10	1251.5	50.5	358	2	S15945
11	1233	49.8	356	2	AH0791
12	1210	48.8	371	2	AC0466
13	1074	43.4	356	2	F71346
14	957	38.6	371	2	H82497
15	425	17.2	1027	2	AC1841
16	415.5	16.8	786	2	D75630
17	383	15.5	383	2	A83603
18	368	14.9	356	2	F87654
19	354	14.3	389	2	AH1937
20	353	14.3	372	2	T45628
21	347.5	14.0	392	2	D96770
22	321.5	13.0	293	2	I40418
23	269.5	10.9	309	2	H89862
24	262.5	10.6	284	2	D84112
25	246.5	10.0	375	2	A83352
26	222	9.0	243	2	E69827
27	181	7.3	247	2	E89956
28	179.5	7.2	239	2	F69961
29	172.5	7.0	583	2	AH1151

30	171	6.9	249	2	G96952	glycerophosphoryl
31	170.5	6.9	583	2	A11510	C-terminal domain
32	168	6.8	232	2	AD1236	glycerophosphodies
33	167	6.7	235	2	A11598	glycerophosphodies
34	165.5	6.7	285	2	E75317	glycerophosphoryl
35	156.5	6.3	240	2	G90544	hypothetical prote
36	154	6.2	353	2	C64187	outer membrane pro
37	152.5	6.2	268	2	AF3245	agrociniopine phosp
38	151.5	6.1	763	2	F96693	hypothetical prote
39	148.5	6.0	271	2	B86918	hypothetical prote
40	147.5	6.0	241	2	S73747	glycerophosphoryl
41	146	5.9	770	2	T04792	hypothetical prote
42	145	5.9	587	2	G89785	hypothetical prote
43	144.5	5.8	244	2	D84232	glycerophosphoryl
44	142.5	5.8	250	2	H83784	glycerophosphodies
45	142	5.7	281	2	H86632	conserved hypothet

ALIGNMENTS

RESULT 1

A43576
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor - Haemophilus influenzae
N:Alternate names: immunoglobulin D-binding protein; protein D
C:Species: Haemophilus influenzae
A:Variety: isolate 772
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 08-Oct-1999
C:Accession: A43576
R:Janson, H.; Heden, L.O.; Grubb, A.; Ruan, M.; Forsgren, A.
Infected. Immun. 59, 119-125, 1991
A:Title: Protein D, an immunoglobulin D-binding protein of Haemophilus influenzae: cloning and characterization of the gene
A:Reference number: A43576; MUID:91099948; PMID:1987023
A:Accession: A43576
A:Molecule type: DNA
A:Residues: 1-364 <JAN>
A:Cross-references: GB:M37487; NID:G148970; PIDN:AAA24998.1; PID:G148971
A:Experimental source: isolate 772
C:Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-364/Product: glycerophosphodiester phosphodiesterase #status predicted <MAT>

Query Match	77.5%	Score	1920	DB	2	Length	364
Best Local Similarity	99.5%	Pred. No.	1.4e-122				
Matches	362	Conservative	0	Mismatches	2	Indels	0
Gaps	0						
QY	1	MDPKTLALSLLAAGVLGCS	SHSSNNANTOMKSDKIIIAHRCASGYLPEHTLES	KALAPA	60		
DB	1	MKLTALSLAAGVLGCS	SHSSNNANTOMKSDKIIIAHRCASGYLPEHTLES	KALAPA	60		
QY	61	QOADYLEQDLAMTKDGR	LVLVHDFLDGLTDVAKKFPFHRHRKDRYYVIDFTLKEIQSLE	120			
DB	61	QOADYLEQDLAMTKDGR	LVLVHDFLDGLTDVAKKFPFHRHRKDRYYVIDFTLKEIQSLE	120			
QY	121	MTENFETKDGKQAQVY	PNRFPMLKSHPRIHTFDEIBIEFIOGLEKSTGKVGVIPEIKAPW	180			
DB	121	MTENFETKDGKQAQV	PNRFPMLKSHPRIHTFDEIBIEFIOGLEKSTGKVGVIPEIKAPW	180			
QY	181	FHHQNGKDIAAETLKV	LYGDKYVDDKTDVYLTQDPNELKRIKTELLPMQMGMDLKLVLQI	240			
DB	181	FHHQNGKDIAAETLKV	LYGDKYVDDKTDVYLTQDPNELKRIKTELLPMQMGMDLKLVLQI	240			
QY	241	AYTDWKETQEKDPKG	YWNVYNDWMFKGMAEVVKYADGVGPGWYMLVNKEESKPDNIV	300			
DB	241	AYTDWKETQEKDPKG	YWNVYNDWMFKGMAEVVKYADGVGPGWYMLVNKEESKPDNIV	300			
QY	301	YTPLVKELAQXNV	VEHPVTVRKDALPEFTDQNMQYDALLNKS	GATGVTFTDPTDGTGEVPL	360		
DB	301	YTPLVKELAQXNV	VEHPVTVRKDALPEFTDQNMQYDALLNKS	GATGVTFTDPTDGTGEVPL	360		
QY	361	KGIK	364				
DB	361	KGIK	364				

Db 121 MTENFETKDGKQAQVYNNRFLWKKSHFRHTPEDEIEFTQGLEKSTGKKVGIYPIKAPW 180
QY 181 FHHQNGKDIAAETLKVLYKGYDKKTDVWYLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKGYDKKTDVWYLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
QY 241 AYTDWKEFQEKDQKPGVWVNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKEFQEKDQKPGVWVNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVFTDPDGTGVBEFL 360
Db 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVFTDPDGTGVBEFL 360
QY 361 KGIK 364
Db 361 KGIK 364

RESULT 5

G64086
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain R
N/Alternate names: Igd-binding protein; protein D
C/Species: Haemophilus influenzae
C/Accession: G64086
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
R/Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kexlavage, R.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Frichman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A./Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A/Reference number: A64000; MUID: 95350630; PMID: 7542800
A/Accession: G64086
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-364 <TIGR>
A/Cross-references: UNIPROT:Q06282; GB:U32751; GB:L42023; NID:g3212200; PIDN:AAC22348.1;
A/Experimental source: strain Rd KW20
C/Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 76.8%; Score 1902; DB 2; Length 364;
Best Local Similarity 98.4%; Pred. No. 2.4e-121;
Matches 358; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDPKTLALSLAAGVLACGSSSHSNMANTQKSKIIIAHRCASGYLPEHTLESKALAPA 60
Db 1 MKLKTALSLAAGVLACGSSSHSNMANTQKSKIIIAHRCASGYLPEHTLESKALAPA 60
QY 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRHDKGRVYVYIDFTLKEIQSLE 120
Db 61 QHSDYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRHDKGRVYVYIDFTLKEIQSLE 120
QY 121 MTENFETKDGKQAQVYNNRFLWKKSHFRHTPEDEIEFTQGLEKSTGKKVGIYPIKAPW 180
Db 121 MTENFETKDGKQAQVYNNRFLWKKSHFRHTPEDEIEFTQGLEKSTGKKVGIYPIKAPW 180
QY 181 FHHQNGKDIAAETLKVLYKGYDKKTDVWYLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKGYDKKTDVWYLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
QY 241 AYTDWKEFQEKDQKPGVWVNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKEFQEKDQKPGVWVNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVFTDPDGTGVBEFL 360
Db 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVFTDPDGTGVBEFL 360
QY 361 KGIK 364
Db 361 KGIK 364

RESULT 6

S59933
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain 6
N/Alternate names: immunoglobulin D-binding protein; protein D
C/Species: Haemophilus influenzae
A/Variety: strain 6-7626
C/Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
C/Accession: S59933; S47335
R/Song, X.M.; Forsgren, A.; Jansson, H.
Infect. Immun. 63, 696-699, 1995
A/Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen
A/Reference number: S59931; MUID: 95122210; PMID: 7822043
A/Accession: S59933
A/Molecule type: DNA
A/Residues: 1-364 <SON>
A/Cross-references: EMBL:Z35658; NID:g525217; PIDN:CAA84717.1; PID:g525218
A/Experimental source: strain 6-7626
C/Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 76.1%; Score 1885; DB 2; Length 364;
Best Local Similarity 97.5%; Pred. No. 3.4e-120;
Matches 355; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPKTLALSLAAGVLACGSSSHSNMANTQKSKIIIAHRCASGYLPEHTLESKALAPA 60
Db 1 MKLKTALSLAAGVLACGSSSHSNMANTQKSKIIIAHRCASGYLPEHTLESKALAPA 60
QY 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRHDKGRVYVYIDFTLKEIQSLE 120
Db 61 QOADYLEODLAWTKDGRVLIHDFLDGLTDVAKKFPFHRHDKGRVYVYIDFTLKEIQSLE 120
QY 121 MTENFETKDGKQAQVYNNRFLWKKSHFRHTPEDEIEFTQGLEKSTGKKVGIYPIKAPW 180
Db 121 MTENFETKDGKQAQVYNNRFLWKKSHFRHTPEDEIEFTQGLEKSTGKKVGIYPIKAPW 180
QY 181 FHHQNGKDIAAETLKVLYKGYDKKTDVWYLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKGYDKKTDVWYLTQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
QY 241 AYTDWKEFQEKDQKPGVWVNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKEFQEKDQKPGVWVNNYNDWMPKPGMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
QY 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVFTDPDGTGVBEFL 360
Db 301 YTPLVKELAQYNVEVHPYTVRKDALPEPFTDVNQMYDALLNKSGATGVFTDPDGTGVBEFL 360
QY 361 KGIK 364
Db 361 KGIE 364

RESULT 7

S59936
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Haemophilus influenzae (strain N
N/Alternate names: Igd-binding protein; protein D
C/Species: Haemophilus influenzae
A/Variety: strain NCTC 8468
C/Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 08-Oct-1999
C/Accession: S59936; S47338
R/Song, X.M.; Forsgren, A.; Jansson, H.
Infect. Immun. 63, 696-699, 1995
A/Title: The gene encoding protein D (hpd) is highly conserved among Haemophilus influen
A/Reference number: S59931; MUID: 95122210; PMID: 7822043
A/Accession: S59936
A/Molecule type: DNA
A/Residues: 1-364 <SON>
A/Cross-references: EMBL:Z35661; NID:g525223; PIDN:CAA84720.1; PID:g525224
A/Experimental source: strain NCTC 8468
C/Keywords: lipoprotein; phosphoric diester hydrolase; surface antigen

Query Match 76.1%; Score 1885; DB 2; Length 364;
Best Local Similarity 97.5%; Pred. No. 3.4e-120;
Matches 355; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTOMKSKIIIIAHRGAGYLPEHTLESKALAPA 60
Db 1 MKLKTALSLLATGTVAGCSSHSSNMANTOMKSKIIIIAHRGAGYLPEHTLESKALAPA 60

Qy 61 QOADYLEODLAWTKDGRVLIWHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIOSLE 120
Db 61 QHADYLEODLAWTKDGRVLIWHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIOSLE 120

Qy 121 MTENFETKDGKQAQVYNNRFPPLWKSHPRIHFFEDIEFIOGLEKSTGKVGIYPEIKAPW 180
Db 121 MTENFETKDGKQAQVYNNRFPPLWKSHPRIHFFEDIEFIOGLEKSTGKVGIYPEIKAPW 180

Qy 181 FHHQNGKDIAAETLKVLYKKGYYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240
Db 181 FHHQNGKDIAAETLKVLYKKGYYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQI 240

Qy 241 AYTDWKETOEKDPKGYWVNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300
Db 241 AYTDWKETOEKDPKGYWVNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIV 300

Qy 301 YTPLVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVFTDFPDGTGVEFL 360
Db 301 YTPLVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVFTDFPDGTGVEFL 360

Qy 361 KGIK 364
Db 361 KGIK 364

RESULT 8
D91019
glycerophosphodiester phosphodiesterase [imported] - Escherichia coli (strain O157:H7, #
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D91019
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; NUID:21156231; PMID:11258796
A:Accession: D91019
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <HAY>
A:Cross-references: UNIPROT:Q8XE17; GB:BA000007; PIDN:BA036547.1; PID:gl3362594; GSPDB:C
A:Experimental source: strain O157:H7, substrain RIND 0509952
C:Genetics:
A:Gene: EC3J24

Query Match 50.9%; Score 1261.5; DB 2; Length 358;
Best Local Similarity 64.7%; Pred. No. 5.3e-78;
Matches 231; Conservative 55; Mismatches 64; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLAGCSSHSSNMANTOMKSKIIIIAHRGAGYLPEHTLESKALAPAQQA 63
Db 6 KTLALSLLAAGVLAGCSSHSSNMANTOMKSKIIIIAHRGAGYLPEHTLPKAMAWAYAQA 59

Qy 64 DYLEDLAWTKDGRVLIWHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIOSLEWTE 123
Db 60 DYLEDLAWTKDGRVLIWHDFHDLGLTDVADRFPPRARKDGRYYAIDFTLDEIKSLKETE 119

Qy 124 NFKETKDGKQAQVYNNRFPPLWKSHPRIHFFEDIEFIOGLEKSTGKVGIYPEIKAPWPHH 183
Db 120 GFDIENGKKVQYVYPCRFPMGKSDFRVHTFEEIEFVQGLNHSHTGKNIIGIYPEIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLYKKGYYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 180 QEGKDIAAETLKVLYKKGYYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 239

Qy 244 DWKETQEKDPKGYWVNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIVVTP 303
Db 240 DWNETQKQKQPGSWVNNYNDWMPKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLWVHPYTVRSKLPYTTVDVNLQYDALLYNKAGVNGLFTDFDPKAVKFL 355

RESULT 10
S15945
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor, periplasmic - Escherich
C:Species: Escherichia coli
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: S15945; S00871; E64994; S72653; S14522
R:Tommassen, J.; Biglmeier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.
Mol. Gen. Genet. 226, 321-327, 1991
A:Title: Characterization of two genes, glpQ and ugqQ, encoding glycerophosphoryl dieste
A:Reference number: S15945; MUID:91238712; PMID:1851953
A:Accession: S15945
A:Molecule type: DNA
A:Residues: 1-358 <TOM>
A:Cross-references: UNIPROT:P09394; EMBL:X56907; NID:G41580; PIDN:CAA40223.1; PID:G41581
R:Biglmeier, K.; Boos, W.; Cole, S.T.

Qy 244 DWKETQEKDPKGYWVNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIVVTP 303
Db 240 DWNETQKQKQPGSWVNNYNDWMPKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLWVHPYTVRSKLPYTTVDVNLQYDALLYNKAGVNGLFTDFDPKAVKFL 355

RESULT 9
F85863
glycerophosphodiester phosphodiesterase [imported] - Escherichia coli (strain O157:H7, #
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F85863
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85863
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STO>
A:Cross-references: UNIPROT:Q8XE17; GB:AB005174; NID:gl2516579; PIDN:AAG57370.1; GSPDB:B
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: glpQ

Query Match 50.9%; Score 1261.5; DB 2; Length 358;
Best Local Similarity 64.7%; Pred. No. 5.3e-78;
Matches 231; Conservative 55; Mismatches 64; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLAGCSSHSSNMANTOMKSKIIIIAHRGAGYLPEHTLESKALAPAQQA 63
Db 6 KTLALSLLAAGVLAGCSSHSSNMANTOMKSKIIIIAHRGAGYLPEHTLPKAMAWAYAQA 59

Qy 64 DYLEDLAWTKDGRVLIWHDFHDLGLTDVAKKFPFHRKDGRIYVDFTLKEIOSLEWTE 123
Db 60 DYLEDLAWTKDGRVLIWHDFHDLGLTDVADRFPPRARKDGRYYAIDFTLDEIKSLKETE 119

Qy 124 NFKETKDGKQAQVYNNRFPPLWKSHPRIHFFEDIEFIOGLEKSTGKVGIYPEIKAPWPHH 183
Db 120 GFDIENGKKVQYVYPCRFPMGKSDFRVHTFEEIEFVQGLNHSHTGKNIIGIYPEIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLYKKGYYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 180 QEGKDIAAETLKVLYKKGYYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 239

Qy 244 DWKETQEKDPKGYWVNNYNDWMPKPGAMAEVVKYADGVGPGWYMLVNKEESKPDNIVVTP 303
Db 240 DWNETQKQKQPGSWVNNYNDWMPKPGAMQVAEYADGIGPDYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQYNVEHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVFTDFPDGTGVEFL 360
Db 299 MVQDAQQNKLWVHPYTVRSKLPYTTVDVNLQYDALLYNKAGVNGLFTDFDPKAVKFL 355

RESULT 10
S15945
glycerophosphodiester phosphodiesterase (EC 3.1.4.46) precursor, periplasmic - Escherich
C:Species: Escherichia coli
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: S15945; S00871; E64994; S72653; S14522
R:Tommassen, J.; Biglmeier, K.; Cole, S.T.; Overduin, P.; Larson, T.J.; Boos, W.
Mol. Gen. Genet. 226, 321-327, 1991
A:Title: Characterization of two genes, glpQ and ugqQ, encoding glycerophosphoryl dieste
A:Reference number: S15945; MUID:91238712; PMID:1851953
A:Accession: S15945
A:Molecule type: DNA
A:Residues: 1-358 <TOM>
A:Cross-references: UNIPROT:P09394; EMBL:X56907; NID:G41580; PIDN:CAA40223.1; PID:G41581
R:Biglmeier, K.; Boos, W.; Cole, S.T.

Mol. Microbiol. 1, 251-258, 1987
A:Title: Nucleotide sequence and transcriptional startpoint of the glpT gene of Escherichia coli
A:Reference number: S00866; MUID:88201663; PMID:3329281
A:Accession: S00871
A:Molecule type: DNA
A:Residues: 1-9 <EIG>
A:Cross-references: EMBL:Y00536; NID:941586; PIDN:CAA68599.1; PID:9581099
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E64994
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-358 <BLAT>
A:Cross-references: GB:AE000314; GB:U00096; NID:gl788570; PIDN:AAC75299.1; PID:gl788572; R:Gonzalez-Gill, G.; Bringham, P.; Kahmann, R.
Mol. Microbiol. 22, 21-29, 1996
A:Title: FIS is a regulator of metabolism in Escherichia coli.
A:Reference number: S72651; MUID:97055418; PMID:8899705
A:Accession: S72653
A:Molecule type: protein
A:Residues: 26, 'I', 28-44 <GON>
C:Genetics:
A:Gene: glpQ
C:Keywords: phosphoric diester hydrolase
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-358/Product: glycerophosphodiester phosphodiesterase #status experimental <MAT>

Query Match 50.5%; Score 1251.5; DB 2; Length 358;
Best Local Similarity 64.1%; Pred. No. 2.5e-77;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;

QY 4 KTLALSLLAAGVLACSSSHSNMANTQMSDKIIIAHAGSGLPEHTLESKALAPAOQA 63
Db 6 KNLSVALMLAGTWIGSGA-----VAEKVVIHAGSGLPEHTLPKAMAYAOQA 59

QY 64 DYLEDQDLAMTKDGRVLVTHDHLGLDTPVAKKPHRRKDGYYVIDFTLKEIQSLEWTE 123
Db 60 DYLEDQDLVMTKDDNLVLDHLDVDRVADFPDRAKDGYYAIDFTLDEIKSLKETE 119

QY 124 NFETKDGKQAQVYNNRFLPKSHRIHTFDEIEFIQGLEKSTGKKGVIYPEIKAPWPHH 183
Db 120 GFDIENGKKVQYTPGRFPMGKSDFRVHTEFEEIEFVQGLNHSTGKNGIYPEIKAPWPHH 179

QY 184 QNGKDIAAETLKVLYKGYDKTDVYLYQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 180 QEGKDIAAETLKVLYKGYDKTDVYLYQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 239

QY 244 DWKETQEKDPKGYVNNYNDWFKPGMAEVVYKADGVPGWYMLVNKEESKPDNIVVTP 303
Db 240 DWNETQQKQPDGWSWYNNYNDWFKPGMAEQVAEYADGIGPDYHMLI-BETSPQGNIKLTG 298

QY 304 LVKELAQYNVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDGTGVEFL 360
Db 299 MVQDAQQNKLVLVHPYTVRSKLPETPDVNLQYDALYNKAGVNGLFTDFPDKAVKFL 355

RESULT 11
AH0791
Glycerophosphodiester phosphodiesterase (EC 3.1.4.46) - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0791
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davies, P.; Davis, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0791
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <PAB>
A:Cross-references: GB:AL513382; PIDN:CAD07514.1; PID:gl6503507; GSPDB:GN00176
C:Genetics:
A:Gene: STY2511
C:Keywords: phosphoric diester hydrolase

Query Match 49.8%; Score 1233; DB 2; Length 356;
Best Local Similarity 64.5%; Pred. No. 4.5e-76;
Matches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2;

QY 4 KTLALSLLAAGVLACSSSHSNMANTQMSDKIIIAHAGSGLPEHTLESKALAPAOQA 63
Db 6 KNLSVALMLAGTWIGSGA-----VAEKVVIHAGSGLPEHTLPKAMAYAOQA 56

QY 64 DYLEDQDLAMTKDGRVLVTHDHLGLDTPVAKKPHRRKDGYYVIDFTLKEIQSLEWTE 123
Db 57 DYLEDQDLVMTKDDNLVLDHLDVDRVADFPDRAKDGYYAIDFTLDEIKSLKETE 116

QY 124 NFETKDGKQAQVYNNRFLPKSHRIHTFDEIEFIQGLEKSTGKKGVIYPEIKAPWPHH 183
Db 117 GFDIENGKKVQYTPGRFPMGKSDFRVHTEFEEIEFVQGLNHSTGKNGIYPEIKAPWPHH 176

QY 184 QNGKDIAAETLKVLYKGYDKTDVYLYQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 243
Db 177 QEGKDIAAETLKVLYKGYDKTDVYLYQTFDFNELKRIKTELLPQMGMDLKLVLQIAYT 236

QY 244 DWKETQEKDPKGYVNNYNDWFKPGMAEVVYKADGVPGWYMLVNKEESKPDNIVVTP 303
Db 237 DWNETQQKQPDGWSWYNNYNDWFKPGMAEQVAEYADGIGPDYHMLV-AEGSTKGNIKLTG 295

QY 304 LVKELAQYNVEVHPYTVRKDALPEFTDVNQMYDALLNKSGATGVFTDPDGTGVEFL 361
Db 296 MVQDAHQNMVHPYTVRADQLPDYATDVNLQYDALYNKAGVNGLFTDFPDKAVMFLQ 353

RESULT 12
AC0466
Glycerophosphodiester phosphodiesterase (EC 3.1.4.46) [imported] - Yersinia pestis (strain
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AC0466
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0466
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <KUR>
A:Cross-references: UNIPROT:Q8ZAH4; GB:AL590842; PIDN:CAC93295.1; PID:gl5981742; GSPDB:G
C:Genetics:
A:Gene: glpQ
C:Keywords: phosphoric diester hydrolase

Query Match 48.8%; Score 1210; DB 2; Length 371;
Best Local Similarity 62.7%; Pred. No. 1.7e-74;
Matches 227; Conservative 54; Mismatches 75; Indels 6; Gaps 4;

QY 4 KTLALSLLAAGVLACSSSHSNMANTQMSDKIIIAHAGSGLPEHTLESKALAPAOQA 60
Db 6 KTLALSLLAASMA-AFAPAAKASADKASAIKVIHAGSGLPEHTLESKALAPAOQA 64

QY 61 QQAQYLEDQDLAMTKDGRVLVTHDHLGLDTPVAKKPHRRKDGYYVIDFTLKEIQSLE 120
Db 65 QQAQYLEDQDLVMTKDDNLVLDHLDVDRVADFPDRAKDGYYAIDFTLDEIKSLK 124

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2005, 11:30:25 ; Search time 171 Seconds
(without alignments)
1389.501 Million cell updates/sec

Title: US-09-719-379A-81
Perfect score: 2477
Sequence: 1 MDPKTLAUSLLAAGVLACCS.....YKFYDNKRIDSTGHHHHH 464

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_prot.*

2: uniprot_treml.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1902	76.8	364	1 GLPQ_HAEIN	Q06282 haemophilus
2	1577	63.7	358	2 Q9CLO4	Q9CLO4 pasteurella
3	1501.5	60.6	368	2 Q65R11	Q65R11 manheimia
4	1498.5	60.5	357	2 Q8RHT6	Q8RHT6 fusobacteri
5	1363	55.0	309	2 Q7P6W6	Q7P6W6 fusobacteri
6	1261.5	50.9	358	2 Q8XE17	Q8XE17 escherichia
7	1254.5	50.6	358	2 Q8CVV6	Q8CVV6 escherichia
8	1251.5	50.5	358	1 GLPQ_ECOLI	P09394 escherichia
9	1251.5	50.5	358	2 Q7UC72	Q7UC72 shigella fl
10	1251.5	50.5	377	2 Q83QU2	Q83QU2 shigella fl
11	1248	50.4	361	2 Q8CZ14	Q8CZ14 erwinia car
12	1235	49.9	356	2 Q8ZNG7	Q8ZNG7 salmonella
13	1233	49.8	356	2 Q8Z556	Q8Z556 salmonella
14	1210.5	48.9	358	2 Q7MZY7	Q7MZY7 photorhabdu
15	1210.5	48.8	371	2 Q66FX7	Q66FX7 yersinia ps
16	1210	48.8	371	2 Q82AH4	Q82AH4 yersinia pe
17	1210	48.8	372	2 Q8D1J5	Q8D1J5 yersinia pe
18	1082	43.7	356	2 Q9X5U6	Q9X5U6 treponema p
19	1074	43.4	356	1 GLPQ_TREPA	Q30405 treponema p
20	1074	43.4	356	2 Q71U57	Q71U57 treponema s
21	1074	43.4	356	2 Q71U58	Q71U58 treponema p
22	1074	43.4	356	2 Q71U59	Q71U59 treponema p
23	1074	43.4	356	2 Q71U64	Q71U64 treponema p
24	1072.5	43.3	341	2 Q7QMYS	Q7QMYS anopheles g
25	1026	41.4	360	2 Q6LT65	Q6LT65 photobacter
26	987.5	39.9	351	2 Q87M77	Q87M77 vibrio para
27	979	39.5	352	2 Q7MDR9	Q7MDR9 vibrio vuln
28	979	39.5	352	2 Q8D6W7	Q8D6W7 vibrio vuln
29	957	38.6	371	2 Q9KN30	Q9KN30 vibrio chol
30	839.5	33.9	338	2 Q9F8D4	Q9F8D4 borrelia tu
31	836	33.8	337	2 Q9F8D3	Q9F8D3 borrelia co

32	830.5	33.5	338	2	Q9F8D5	Q9F8D5 borrelia pa
33	825.5	33.3	334	2	Q6UKS8	Q6UKS8 borrelia mi
34	821.5	33.2	340	2	Q45201	Q45201 borrelia he
35	811.5	32.8	336	2	Q6UKS9	Q6UKS9 borrelia lo
36	803.5	32.4	333	2	Q9F8D6	Q9F8D6 borrelia cr
37	797.5	32.2	333	2	Q9ETK6	Q9ETK6 borrelia re
38	428.5	17.3	372	2	Q7YZB5	Q7YZB5 bombyx mori
39	425	17.2	1027	2	Q8Z028	Q8Z028 anabaena sp
40	415.5	16.8	786	2	Q9RZL4	Q9RZL4 deinococcus
41	385	15.5	392	2	Q8CJ18	Q8CJ18 streptomyce
42	383	15.5	383	2	Q916E6	Q916E6 pseudomonas
43	368	14.9	356	2	Q9A3D2	Q9A3D2 caulobacter
44	360	14.5	393	2	Q827Y5	Q827Y5 streptomyce
45	354	14.3	389	2	Q8Y104	Q8Y104 anabaena sp

ALIGNMENTS

RESULT 1

ID	GLPQ_HAEIN	STANDARD;	PRT;	364 AA.
AC	Q06282;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Glycerophosphoryl diester phosphodiesterase precursor (EC 3.1.4.46)			
DE	(Glycerophosphodiester phosphodiesterase) (Surface-exposed lipoprotein			
DE	D) (Protein D) (Immunoglobulin D-binding protein) (IgD-binding			
DE	protein).			
GN	Name=glpQ; Synonyms=hpd; OrderedLocusNames=HI0689;			
OS	Haemophilus influenzae.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Haemophilus.			
OX	NCBI_TaxID=727;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RD / KW20 / ATCC 51907;			
RX	MEDLINE=95350630; PubMed=7542800;			
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,			
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,			
RA	McKenny K., Sutton G.G., FitzHugh W., Fields C.A., Gocayne J.D.,			
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,			
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,			
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,			
RA	Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,			
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,			
RA	Venter J.C.;			
RT	"Whole-genome random sequencing and assembly of Haemophilus influenzae			
RL	Rd.";			
RL	Science 269:496-512(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NTHi 772;			
RX	MEDLINE=91099948; PubMed=1987023;			
RA	Janson H., Heden L.-O., Grubb A., Ruan M., Forsgren A.;			
RT	"Protein D, an immunoglobulin D-binding protein of Haemophilus			
RT	influenzae: cloning, nucleotide sequence, and expression in			
RT	Escherichia coli";			
RL	Infect. Immun. 59:119-125(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Minna / Serotype B;			
RX	MEDLINE=94011360; PubMed=8104899;			
RA	Janson H., Ruan M., Forsgren A.;			
RT	"Limited diversity of the protein D gene (hpd) among encapsulated and			
RT	nonencapsulated Haemophilus influenzae strains.";			
RL	Infect. Immun. 61:4546-4552(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=3639, 3640, 6-7626, Eagan / Serotype B, HK695 / Serotype B, and			
RX	NCTC 8468 / Serotype B;			
RX	MEDLINE=95122210; PubMed=7822043;			

RA Song X.-M., Forsgren A., Janson H.;
 RT "The gene encoding protein D (hpd) is highly conserved among
 RT Haemophilus influenzae type b and nontypeable strains.";
 RL Infect. Immun. 63:696-699(1995).
 [5]
 RP CHARACTERIZATION.
 RC STRAIN=NTH1 772;
 RX MEDLINE=92192801; PubMed=1548059;
 RA Janson H., Heden L.-O., Forsgren A.;
 RT "Protein D, the immunoglobulin D-binding protein of Haemophilus
 RT influenzae, is a lipoprotein.";
 RL Infect. Immun. 60:1336-1342(1992).
 CC -!- FUNCTION: Glycero-phosphoryl diester phosphodiesterase hydrolyzes
 CC deacylated phospholipids to G3P and the corresponding alcohols.
 CC Has a specific affinity for human immunoglobulin D myeloma
 CC protein.
 CC -!- CATALYTIC ACTIVITY: A glycerophosphodiester + H(2)O = an alcohol +
 CC sn-glycerol 3-phosphate.
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor.
 CC -!- PTM: Contains both ester- and amide-linked fatty acids.
 CC -!- MISCELLANEOUS: The sequence shown is that of strains NTH1 772 and
 CC RD / KW20.
 CC -!- SIMILARITY: Belongs to the glycerophosphoryl diester
 CC phosphodiesterase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 DR EMBL; U32751; AAC22348.1; -;
 DR EMBL; M37487; AAA24998.1; -;
 DR EMBL; L12445; AAA24999.1; -;
 DR EMBL; Z35656; CAA84715.1; -;
 DR EMBL; Z35657; CAA84716.1; -;
 DR EMBL; Z35658; CAA84717.1; -;
 DR EMBL; Z35659; CAA84718.1; -;
 DR EMBL; Z35660; CAA84719.1; -;
 DR EMBL; Z35661; CAA84720.1; -;
 DR PIR; G64086; G64086;
 DR PIR; S59934; S59934.
 DR TIGR; H10689; -;
 DR InterPro; IPR004129; GPPD.
 DR Pfam; PF03009; GPPD; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Complete proteome: Glycerol metabolism; Hydrolase; Lipoprotein;
 KW Outer membrane; Palmitate; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 364
 FT FT
 FT FT
 FT LIPID 19 19
 FT LIPID 19 19
 FT LIPID 19 19
 FT VARIANT 13 13
 FT VARIANT 16 16
 FT VARIANT 25 25
 FT VARIANT 28 28
 FT VARIANT 34 34
 FT VARIANT 62 62
 FT VARIANT 63 63
 FT VARIANT 98 98
 FT VARIANT 99 99
 FT VARIANT 144 144
 FT VARIANT 168 168

FT VARIANT 191 191
 FT VARIANT 253 253
 FT VARIANT 310 310
 FT VARIANT 327 327
 FT VARIANT 338 338
 FT VARIANT 364 364
 FT VARIANT 364 364
 SQ SEQUENCE 364 AA; 41902 MW; A6079B3ABF70E820 CRC64;
 Query Match 76.8%; Score 1902; DB 1; Length 364;
 Best Local Similarity 98.4%; Pred. No. 5.8e-124;
 Matches 358; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 MDPKTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
 Db 1 MKLKTALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPA 60
 Qy 61 QQADYLEQDLAMTKDGLVVIHDFLDGLTDVAKKFFPHRHKDGRYYVIDFTLKEIOSLE 120
 Db 61 QHSDYLEQDLAMTKDGLVVIHDFLDGLTDVAKKFFPHRHKDGRYYVIDFTLKEIOSLE 120
 Qy 121 MTENFETKDGKQAVYNNRFPPLWKSHPRIHTFEDEIEFIOGLEKSTCKVGIYPEIKAPW 180
 Db 121 MTENFETKDGKQAVYNNRFPPLWKSHPRIHTFEDEIEFIOGLEKSTCKVGIYPEIKAPW 180
 Qy 181 FHHQNGKDIAAETLKVLLKKGDKYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVL 240
 Db 181 FHHQNGKDIATETLKVLLKKGDKYDKTDMVYLQTFDFNELKRIKTELLPQMGMDLKLVL 240
 Qy 241 AYTDWKETQEKDPKGYWVNTNYDMFKPGMAEYVYADGVGPGWYMLVNKESKPDNIV 300
 Db 241 AYTDWKETQEKDPKGYWVNTNYDMFKPGMAEYVYADGVGPGWYMLVNKESKPDNIV 300
 Qy 301 YTPLVKELAQYVNEVHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVTFDPDTGVVEFL 360
 Db 301 YTPLVKELAQYVNEVHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVTFDPDTGVVEFL 360
 Qy 361 KGIK 364
 Db 361 KGIK 364
 RESULT 2
 Q9CL04 PRELIMINARY; PRT; 358 AA.
 AC O9CL04; 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE GlpQ.
 GN Name=g1pQ; OrderedLocusNames=PM1444;
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 RN [1]
 RN NCBI_TaxID=747;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=X-73;
 RX PubMed=14738894; DOI=10.1016/j.micinf.2003.09.023;
 RA Lo M., Boyce J.D., Wilkie I.W., Adler B.;
 RT "Characterization of two lipoproteins in Pasteurella multocida.";
 RL Microbes Infect. 6:58-67(2004).

Df 299 MVQDAQNKLVVHPHYTVRSKPLPYTTDDVNLQYDALYNKAGVNGLFTDFFDPDKAVKFL 355

RESULT 7
QCWV6 PRELIMINARY; PRT; 358 AA.
ID QBCV6 AC QBCV6 ID QBCV6 AC QBCV6
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBurel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBurel. 24, Last annotation update)
DE Glycerophosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46).
GN Name:glo; OrderedLocusNames=c2780;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Escherichia.
NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.J., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.P., Rose D.B., Zhou S., Schwartz D.C., Ferna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DL EMBL: AE016763; AN81234.1; -
DR GO: GO:0008889; F:glycerophosphodiester phosphodiesterase act.. ; IEA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR GO: GO:0006071; P:glycerol metabolism; IEA.
DR InterPro: IPR004129; GDPD.
DR Pfam: PF03009; GDPD; 1.
KW Complete proteome; Hydrolase.
SQ *SEQUENCE 358 AA; 40826 MW; C633D35EA86591F8 CRC64;

Query Match 50.6%; Score 1254.5; DB 2; Length 358;
Best Local Similarity 64.4%; Pred.No.5.6e-79;
Matches 230; Conservative 55; Mismatches 65; Indels 7; Gaps 2;

Oy 4 KTLALSLLAGVLGACSSSHSNANTOMKSDKIILAHRGASGYLPHTLSKKALAFQAQ 63
Db : :::: |
6 KNLSMAIMSTMIVMGSSAMAAD-----SNEKIVIAHRGASGYLPHTLPKAMAYAQA 59

Oy 64 DYLEODLAMTKOGELVVIHDFLDGLTDVAKKPHRHKDGRYVIDFTLKETQSLEMTE 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|
60 DYLEDQLVMTKDHLVVLVDHYLDLVADVADRFPDRARKDGRYYAIDFTLTDSKLFPE 119

Oy 124 NFETKDQGQQVYPNRRPPLWKSHPRIHTFDSTEFIOGLEKSCKGVGIYPEIKAPWFHH 183
Db : |
120 GFDIENGKVQTVPGRPFMKGSDRFVTTFEEIEFVGQLNHSGNKIGIYPEIKAPWFHH 179

Oy 184 QNGKDIAAEITLVLYKYGDYDKTDMVYLQTFDNELKRITELLPMQMGMALKVLQIIAYT 243
Db : |
180 QEKGDIAAKTLEVLIKYGTYGTGGDYLCFCFADELKRIKNVLEPKMGMDLNVLQIIAYT 239

Oy 244 DWKETQEKPFGYWNNNYNDMPKPGAMEVKYADGVCGFWMLVNKESKPDNIYTP 303
Db : |
240 DMNETOOQPQDGWNWNNYNDMPKPGAMKQVAEYADGIPGDYHMLI-EETSQFNKITLG 298

Oy 304 LVKEAQNVVEVHPHYTVRKDALPEFFTDDNQMYDALLNKGSGATGVFTDPDTGEVEL 360
Db : |
299 MVQDAQNKLVVHPHYTVRSKPLPYTTDDVNLQYDALYNKAGVNGLFTDFFDPDKAVKFL 355

Result 8
GLPO_ECOLI STANDARD; PRT; 358 AA.
ID GLPO_ECOLI AC P9334;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)

or send an email to license@iab-sib.ch).

```
CC -----
CC EMBL; X56907; CAA40323.1; -.
DR EMBL; U00096; AAC75299.1; -.
DR EMBL; D90855; BAA16058.1; -.
DR EMBL; Y00536; CAA68599.1; -.
DR PIR; S15945; S15945.
DR ECODBASE; C039.0; 6TH EDITION.
DR ECHOBASE; EB0394; -.
DR Ecogene; EG10399; gipQ.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome; Direct protein sequencing; Glycerol metabolism;
KW Hydrolase; Multigene family; Periplasmic; Signal.
FT SIGNAL 1 25
FT CHAIN 26 358
FT Glycerophosphoryl diester
FT phosphodiesterase, periplasmic.
FT D -> I (in Ref. 5).
FT CONFLICT 27
FT SEQUENCE 358 AA; 40843 MW; B08FD27399641616 CRC64;

Query Match 50.5%; Score 1251.5; DB 1; Length 358;
Best Local Similarity 64.1%; Pred. No. 9.1e-79;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLGACSSHSSNMANTOMKSKIIIAHAGSGYLPPEHTLESKALAFQAQA 63
Db 6 KNLSSAIMSTIVMGSSAMAAD-----SNEKIVIAHAGSGYLPPEHTLPKAMAYAQQA 59

Qy 64 DYLEQDLAMTKDGRVLVHDFHLDGLTDVAKKFPFHRKDGYYVDFTLKEIOSLEWTE 123
Db 60 DYLEQDLVMTKDDNLVVLHDLVDRVTDVADRFDPDRARKDGRYYAIDFTLDEIKSLKTE 119

Qy 124 NFETKDGKQAOQVYVNPFRPLMKSHPRHTFEDEIFEOGLEKSTGKKGVIYPIKAPWPHH 183
Db 120 GFDIENKGVQVTPGRFPMGKSDFRVHTFEEIEFVQGLNHSSTGKNGIYPIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLLKKGVDKKTDMVYLQTFDFNELKRIKTELLPQMGMIDLKVLQIAYT 243
Db 180 QEGKDIAAKTLEVLKKGYYTGKDKVKYLQCFDADELKRIKNELEPKMGMLNLVQLIAYT 239

Qy 244 DWKETQEKDPKGYWVNYNDWMFKPGMAEVVYKADGVGPGWYMLVKNKESKPDNIYVTP 303
Db 240 DWNETQKQPDGWSWVNYNDWMFKPGMAKQVAEYADGIGDPYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQNVNVEHPYTVRKDALPEFTVDNQMDALLNKSGATGVTDFDPDTGVEFL 360
Db 299 MVQDAQQNKLIVVHPYTVRSKPLPEYTPDNLQYDALYNKAGVNGLFTDFDPDKAVKFL 355

RESULT 9
Q7UC72 PRELIMINARY; PRT; 358 AA.
AC Q7UC72;
AT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycerophosphodiester phosphodiesterase.
GN Name=gipQ; OrderedLocuNames=S2454;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RT
```

```
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AE016986; AAP17654.1; -.
DR GO; GO:0008899; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
SQ SEQUENCE 358 AA; 40893 MW; ABCD1B395FE41D41 CRC64;

Query Match 50.5%; Score 1251.5; DB 2; Length 358;
Best Local Similarity 64.1%; Pred. No. 9.1e-79;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;

Qy 4 KTLALSLLAAGVLGACSSHSSNMANTOMKSKIIIAHAGSGYLPPEHTLESKALAFQAQA 63
Db 6 KNLSSAIMSTIVMGSSAMAAD-----SNEKIVIAHAGSGYLPPEHTLPKAMAYAQQA 59

Qy 64 DYLEQDLAMTKDGRVLVHDFHLDGLTDVAKKFPFHRKDGYYVDFTLKEIOSLEWTE 123
Db 60 DYLEQDLVMTKDDHLVVLHDLVDRVTDVADRFDPDRARKDGRYYAIDFTLDEIKSLKTE 119

Qy 124 NFETKDGKQAOQVYVNPFRPLMKSHPRHTFEDEIFEOGLEKSTGKKGVIYPIKAPWPHH 183
Db 120 GFDIENKGVQVTPGRFPMGKSDFRVHTFEEIEFVQGLNHSSTGKNGIYPIKAPWPHH 179

Qy 184 QNGKDIAAETLKVLLKKGVDKKTDMVYLQTFDFNELKRIKTELLPQMGMIDLKVLQIAYT 243
Db 180 QEGKDIAAKTLEVLKKGYYTGKDKVKYLQCFDADELKRIKNELEPKMGMLNLVQLIAYT 239

Qy 244 DWKETQEKDPKGYWVNYNDWMFKPGMAEVVYKADGVGPGWYMLVKNKESKPDNIYVTP 303
Db 240 DWNETQKQPDGWSWVNYNDWMFKPGMAKQVAEYADGIGDPYHMLI-EETSQPGNIKLTG 298

Qy 304 LVKELAQNVNVEHPYTVRKDALPEFTVDNQMDALLNKSGATGVTDFDPDTGVEFL 360
Db 299 MVQDAQQNKLIVVHPYTVRSKPLPEYTPDNLQYDALYNKAGVNGLFTDFDPDKAVKFL 355

RESULT 10
Q83QU2 PRELIMINARY; PRT; 377 AA.
AC Q83QU2;
AT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycerophosphodiester phosphodiesterase, periplasmic.
GN Name=gipQ; OrderedLocuNames=SF321;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
DR EMBL; AE015248; AAN43836.1; -.
DR GO; GO:0008899; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome.
SQ SEQUENCE 377 AA; 42920 MW; B4EA7855FACAA936 CRC64;

Query Match 50.5%; Score 1251.5; DB 2; Length 377;
Best Local Similarity 64.1%; Pred. No. 9.7e-79;
Matches 229; Conservative 56; Mismatches 65; Indels 7; Gaps 2;
```

```

QY 4 KTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPAOQA 63
Db 25 KNLMSAIMSTIVMGSSMAAD-----SNEKIVIAHRGASGYLPEHTLPKAMAWAQA 78
QY 64 DYLEDQDLAMTKDGLRVLVHDFHDLGLTDVAKKFPFHRKDGRRYVYDFTLKEIQSLEWTE 123
Db 79 DYLEDQDLVMTKDDHLVLDHLDLDRVTDVADFPDRKDGRRYVYDFTLDEIKSLKTE 138
QY 124 NFETKDGKQAOQVYNNRPLPLKWSHFRHTFEDEIEFIQGLEKSTGKKGVIYPIKAPWFFH 183
Db 139 GFDIENGKKVQYVYPCRFPMGKSDFRVHTTPEEBIEFVQGLNHSFGKNIYPIKAPWFFH 198
QY 184 QNGKDIAETLKVLYKGYDKTDMVYLTQTFDNLKRIKTELLPQMGMDLKLVLQIAYTD 243
Db 199 QGKDIAETLKVLYKGYDKTDMVYLTQTFDNLKRIKTELLPQMGMDLKLVLQIAYTD 258
QY 244 DNKETQEKDPKGYWVNNYNDMMFKPGMAEVVKYADGVPGMVYLVNKEESKPDNIVYTP 303
Db 259 DNETQOQKDPGSGWVNNYNDMMFKPGMAEVVKYADGVPGMVYLVNKEESKPDNIVYTP 317
QY 304 LVKELAQYNVEVHPYTVRKDALPEFTDVNMYDALLNKSGATGVFTDPDTPGVFL 360
Db 318 MVQDAQONKLVVHPYTVRSKLPETVDPVNLQYDLYNKAGVNGLFTDPDPAKAVFL 374

RESULT 11
Q6CZ14
ID Q6CZ14 PRELIMINARY; PRT; 361 AA.
AC Q6CZ14;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DE Glycero-phosphoryl diester phosphodiesterase, periplasmic (EC 3.1.4.46).
GN Name=glpQ; OrderedLocustNames=ECA4167;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia N., Pritchard L., Holden M.T.G., Hymn L.J.,
RA Holve M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagsels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmund G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RA "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX50851; CAG77064.1; -.
DR GO; GO:0008889; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 361 AA; 41551 MW; EB3E20B4555AB5F6 CRC64;

Query Match 50.4%; Score 1248; DB 2; Length 361;
Best Local Similarity 63.9%; Pred. No. 1.6e-78;
Matches 228; Conservative 55; Mismatches 70; Indels 4; Gaps 2;

QY 5 TLLSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPAOQA 64
Db 6 TSLISALSVSFSSVSSNSVSASS---TDKIVIAHRGASGYLPEHTLPKAMAWAQAQVD 62
QY 65 YLEQDLAMTKDGLRVLVHDFHDLGLTDVAKKFPFHRKDGRRYVYDFTLKEIQSLEWTE 124

```

```

Db 63 YLEQDLVITKONALIVLHDLDRVTDVAERFPQARQKGRFYAIDFTLKEIKSLKFTSG 122
QY 125 FETKDGKQAOQVYNNRPLPLKWSHFRHTFEDEIEFIQGLEKSTGKKGVIYPIKAPWFFH 184
Db 123 PDKNGKQVQSVYNNRPLPLKWSHFRHTTQEBIEFIQGLNHSFGKNIYPIKAPWFFH 182
QY 185 NKGDIATETLKVLYKGYDKTDMVYLTQTFDNLKRIKTELLPQMGMDLKLVLQIAYTD 244
Db 183 EGKDIATETLKVLYKGYDKTDMVYLTQTFDNLKRIKTELLPQMGMDLKLVLQIAYTD 242
QY 245 WKETQEKDPKGYWVNNYNDMMFKPGMAEVVKYADGVPGMVYLVNKEESKPDNIVYTP 304
Db 243 WHETQEKADGSWVNNYNDMMFKPGMAEVVKYADGVPGMVYLVNKEESKPDNIVYTP 301
QY 305 VKELAQYNVEVHPYTVRKDALPEFTDVNMYDALLNKSGATGVFTDPDTPGVFL 361
Db 302 VKEAHQHLKQVHPYTVRDLPTAKYDQVLYDLYHQAQVDGLFTDPDPAKAVFL 358

RESULT 12
Q8ZNG7
ID Q8ZNG7 PRELIMINARY; PRT; 356 AA.
AC Q8ZNG7;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Periplasmic glycerophosphodiester phosphodiesterase (EC 3.1.4.46).
GN Name=glpQ; OrderedLocustNames=STM2282;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RC MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
DR EMBL; AE008802; AAL21183.1; -.
DR GO; GO:0008889; P:glycerophosphodiester phosphodiesterase act. .; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 356 AA; 40425 MW; 527261A667E09C44 CRC64;

Query Match 49.9%; Score 1235; DB 2; Length 356;
Best Local Similarity 64.5%; Pred. No. 1.3e-77;
Matches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2;

QY 4 KTLALSLLAAGVLAGCSSHSSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAPAOQA 63
Db 6 KNLVSALMLAGMTGSGA-----VAAEKVIAHRGASGYLPEHTLPKAMAWAQA 56
QY 64 DYLEDQDLAMTKDGLRVLVHDFHDLGLTDVAKKFPFHRKDGRRYVYDFTLKEIQSLEWTE 123
Db 57 DYLEDQDLVMTKDDHLVLDHLDLDRVTDVADFPDRKDGRRYVYDFTLDEIKSLKTE 116
QY 124 NFETKDGKQAOQVYNNRPLPLKWSHFRHTFEDEIEFIQGLEKSTGKKGVIYPIKAPWFFH 183
Db 117 GFDIENGKKVQYVYPCRFPMGKSDFRVHTTPEEBIEFVQGLNHSFGKNIYPIKAPWFFH 176
QY 184 QNGKDIAETLKVLYKGYDKTDMVYLTQTFDNLKRIKTELLPQMGMDLKLVLQIAYTD 243
Db 177 QGKDIAETLKVLYKGYDKTDMVYLTQTFDNLKRIKTELLPQMGMDLKLVLQIAYTD 236

```



```

Qy 244 DWKETQEKDPKGYWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 303
Db 237 DWNETQKQDPGRWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 295
Qy 304 LVKELAQYVNEVHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVTDFPDPTGVEFLK 361
Db 296 MVQDAHQNMVWHPYTVRADQLPDYATDVNQYDILYNKAGVDGLTFDFPDKAVMFLQ 353

RESULT 13
Q8Z556 PRELIMINARY; PRT; 356 AA.
ID Q8Z556 Q8Z556; Q7CB71;
AC Q8Z556; Q7CB71;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Glycero-phosphoryl diester phosphodiesterase periplasmic (EC 3.1.4.46).
GN Name=glpQ; OrderedLocusNames=STY2511, t0582;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin P., Hogue A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
RL EMBL; AL627274; CAD07514.1; -.
DR EMBL; AE016836; AA068288.1; -.
DR GO; GO:0008889; F:glycerophosphodiester phosphodiesterase act. . .; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 356 AA; 40437 MW; 51AF570169108CAA CRC64;

Query Match 49.8%; Score 1233; DB 2; Length 356;
Best Local Similarity 64.5%; Pred. No. 1.7e-77;
Matches 231; Conservative 47; Mismatches 70; Indels 10; Gaps 2;

Qy 4 KTLALSLLAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQA 63
Db 6 KNLSSVALLAGMTGTGSA-----VAEKVIAHRGASGYLPEHTLPKAMAYAQGA 56

Qy 64 DYLEDQLANTKDGRLVVIHDFLDGLTDVAKKFPHRHRKDGYYVDFTLKEIOSLEWTE 123
Db 57 DYLEDQLVMTKODHLVLDHLDVTDVADFPDRARCKGYYAIDFTLDEIKSLKFT 116

Qy 124 NFETKDGKQAQVYVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 183
Db 117 GFDIENGKKVQTFGRFPMGKSDPRIHTFEIEFVQGLNHSFGKNGIYPIKAPWPHH 176

Qy 124 NFETKDGKQAQVYVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 303
Db 236 DWNETQKQDPGRWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 294
Qy 304 LVKELAQYVNEVHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVTDFPDPTGVEFLK 361

```

```

Qy 184 QNGKIDIAAETLKVLYKGYDKYKTDVMVYLQTFDFNFELKRIKTELLPQGMGLKLVQLIAYT 243
Db 177 QEGKIDIAAETLKVLYKGYDKYKTDVMVYLQTFDFVAELKRIKNELEPKMGMDLNLVQLIAYT 236
Qy 244 DWKETQEKDPKGYWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 303
Db 237 DWNETQKQDPGRWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 295
Qy 304 LVKELAQYVNEVHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVTDFPDPTGVEFLK 361
Db 296 MVQDAHQNMVWHPYTVRADQLPDYATDVNQYDILYNKAGVDGLTFDFPDKAVMFLQ 353

RESULT 14
Q7MZY7 PRELIMINARY; PRT; 358 AA.
ID Q7MZY7 Q7MZY7;
AC Q7MZY7; Q7MZY7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycero-phosphoryl diester phosphodiesterase, periplasmic
DE (Glycerophosphodiester phosphodiesterase).
GN Name=glpQ; OrderedLocusNames=plu4120;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Bouraux-Eude C., Chandler M., Charles J.-P.,
RA Daessle E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Dassa A.C., Lanois A., Powell K., Siguer P., Gancin R., Wingard V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
RL EMBL; BX571872; CAS16492.1; -.
DR Photoblast; plu4120; -.
DR GO; GO:0008889; F:glycerophosphodiester phosphodiesterase act. . .; IEA.
DR GO; GO:0006071; P:glycerol metabolism; IEA.
DR InterPro; IPR004129; GDPD.
DR Pfam; PF03009; GDPD; 1.
KW Complete proteome.
SQ SEQUENCE 358 AA; 41182 MW; 0BBB12D55785A8C5 CRC64;

Query Match 48.9%; Score 1210.5; DB 2; Length 358;
Best Local Similarity 62.0%; Pred. No. 6.4e-76;
Matches 222; Conservative 54; Mismatches 71; Indels 11; Gaps 2;

Qy 4 KTLALSLLAGVLACGSSSHSNMANTQMSDKIIIAHRGASGYLPEHTLESKALAFQAQA 63
Db 6 KTWIIGIIL-----TSSMSGIAQAADKIVIAHRGASGYLPEHTLPKAMAYAQGA 55

Qy 64 DYLEDQLANTKDGRLVVIHDFLDGLTDVAKKFPHRHRKDGYYVDFTLKEIOSLEWTE 123
Db 56 DYLEDQLVMTKODHLVLDHLDVTDVANKFPNARQDGRYYAIDFTLSEIKSLKFT 115

Qy 124 NFETKDGKQAQVYVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 183
Db 116 GFDIKNDQIQNFSNRPPIWKSDFRIHTFEEIEFVQGLNHSFGKNGIYPIKAPWPHH 175

Qy 184 QNGKIDIAAETLKVLYKGYDKYKTDVMVYLQTFDFNFELKRIKTELLPQGMGLKLVQLIAYT 243
Db 176 KEGKIDISTKVLAVLKGYYTKSKDYLYLQCFDPTNKLKRIKNELEPKMGMDLNLVQLIAYT 235
Qy 244 DWKETQEKDPKGYWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 303
Db 236 DWNETQKQDPGRWVNNYDMWFKPGMAEYVVKYADGVGPGWYMLVNKESKPDNIIVTP 294
Qy 304 LVKELAQYVNEVHPYTVRKDALPEFFTDVNMQYDALLNKSGATGVTDFPDPTGVEFLK 361

```


Search completed: June 22, 2005, 11:40:43
Job time : 175 secs

THIS PAGE BLANK (USPTO)